DOCUMENT RESUME

ED 197 423 EA 013 177

AUTHOR Licata, Joseph W.
TITLE Systemic Appraisal of Educational Leadership Personnel. Informal Publications.
INSTITUTION Ohio State Univ., Columbus. Mershon Center.
PUB DATE 80
NOTE 75p.

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Administrator Evaluation; *Job Performance; Management Development; Principals; *Professional Development; Self Evaluation (Individuals)

ABSTRACT

This text presents a model that systemically relates professional development and evaluation in an appraisal of educational leadership personnel. Developmentally, appraisal is seen as a diagnostic, relatively non-judgmental system by which educational administrators gather data about their professional performance and growth. Evaluatively, appraisal is seen as a data gathering system by which the organization can make necessarily judgmental decisions about its educational leadership personnel. While both development and evaluation use teacher assessments of the administrator as a foundation, the two procedures are distinct and proceed independently. The development process, which remains confidential, begins with the administrator assessing the teacher data. The administrator initiates an external observation of himself or herself by two trained individuals. The observation is followed by a developmental conference designed to create a personal program for professional growth. The evaluation process involves an assessment of the teacher data by a supervisor who initiates evaluation. Two trained individuals observe the administrator's performance, leading to an evaluation conference which offers recommendations and recognition. Implementation of systemic appraisal involves five steps: (1) decide to implement the system; (2) develop a consensus definition of performance; (3) train system participants; (4) field test the system; (5) evaluate the field test results. (JK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED197473

INFORMAL PUBLICATIONS

SYSTEMIC APPRAISAL OF EDUCATIONAL LEADERSHIP PERSONNEL

by

Joseph W. Licata

MERSHON CENTER

Programs of Research and Education in National Security, Leadership, and Public Policy

The Ohio State University
199 West 10th Avenue
Columbus, Ohio 43201

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
J. W. Licata
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Preface

The following text was made possible by support from the Mershon Center for Research and Education in National Security and Policy Sciences, The Ohio State University and the Department of Educational Administration, College of Education, The Ohio State University. While the thoughts and opinions expressed in the text are the sole responsibility of its author and do not necessarily reflect the policies of either supporting agency, this author would like to express specific thanks for the encouragement and help of Vern Cunningham, Margaret Hermann, Charles Hermann and Lonnie Wagstaff. This paper reflects in part the on-going efforts of both agencies with regard to research and development about leadership in the policy sciences. The purpose of the paper is to sketch out an idea that may have application in the development of leadership personnel. The presentation does not present a definitive operational product, but a guide to be creatively adapted in the field to improve leadership in education and other arenas.

Systemic Appraisal of Educational Leadership Personnel

In educational organizations the process of professional appraisal can be conceptually defined to include the generation of data for the development and evaluation of performance. Appraisal can produce diagnostic data for personal and professional development and evaluative data for the organization's evaluation of personnel. Operationally, the process of appraisal in educational organizations is often much less than this conceptual definition. Evaluations of professional performance, resulting in an annual conference between supervisor and educator, have been characterized by some as no more than ritual or "ceremonial congratulation".[1] Professional development, without a definitive and on-going diagnostic data base, is often confined to brief "suggestions for improvement" by the supervisor and "one-shot," in-service meetings at the end of the work day to fatigued professionals concentrating more on the clock than the speaker. In many quarters, these processes have such tarnished reputations that many professionals doubt the possibility that they can ever be successfully implemented in educational organizations.

Rather than abandoning professional appraisal and segmenting the key processes of professional development and evaluation, there is a need to organize appraisal in ways that systemically relate data collection for development to data collection for evaluation. If we define development as a way of helping professional educators accomplish their organizational role expectations and evaluation as a way of determining the extent to which role expectations have been met - to which goals have been attained, then professional appraisal makes sense only when it is organized and operationalized to provide

for systemic integration of these two dimensions. In other words, appraisal helps relate organizational members to organizational processes and goals as a means toward enhanced organizational goal attainment.

Appraising Educational Leadership

As defined, the processes of appraisal can be viewed in terms of evaluation and development of professional performance. In a contemporary sense, these dimensions of appraisal in educational organizations are primarily operationalized in terms of the systematic observation and development of teaching performance. Ironically, while educational administrators are usually responsible for the appraisal of teaching, they rarely receive such feedback themselves. Administering an educational organization can be a challenging undertaking, yet little is being done to appraise and develop the professional performance of these educational leaders. In a recent national survey of secondary school principals, while most principals reported having formally organized evaluation programs for their teachers, few reported such programs in operation relevant to their own evaluation or development.[2]

When it does occur, the appraisal of educational leadership personnel usually focuses on the evaluation of their professional performance for the purpose of making personnel decisions relating to salary, tenure, promotions, and in rare cases, demotion or dismissal. With some exceptions,[3] the use of appraisal processes to guide the professional development of educational administrators has been rare, viewed as a possible serendipitous residual of an occasional evaluation conference. While new appraisal patterns in the professional development and evaluation of teaching stimulate on-going inquiry and experimentation,[4] there is clearly a need to consider such experimentation

in expanding the operational meaning of appraisal for educational leadership personnel.

The purpose of this text is to present a model for the appraisal of educational leadership personnel which systemically relates professional development and evaluation. The model primarily features the use of systematic data collection and analysis as a means by which rational decisions can be made with respect to professional development and evaluation. While the model will be specified to school administrators for the purpose of giving specific examples in description, there is no intent to suggest its exclusive use with these professionals only. Certainly, educational leadership personnel in various organizations may wish to experiment with the model in improving their systems of professional appraisal. In presenting such a model, called here "systemic appraisal," the intent is to stimulate its field testing as an alternative pattern of using appraisal to accomplish improved professional development and evaluation.

Need for Systemic Appraisal

Tensions arising from an apparent incongruence between evaluation and professional development processes have often been unwanted and unintended outcomes of appraisal patterns in education. While both appraisal for evaluation and professional development have much in common with respect to desired outcomes, it is often difficult to accomplish both without threatening the security or autonomy of the professional. For instance, let's consider an annual evaluation conference between a superintendent of schools and a school principal. Both participants know that the purpose of the conference is to articulate or develop a general valuing of the principal's professional

performance on a good-bad continuum and to use that valuing in making a relevant personnel decision, e.g., salary changes, organizational assignments or tenure. The superintendent may attempt to play down this emphasis by suggesting that an important purpose of the conference is to provide feedback to the principal which might be helpful in self-development. Obviously, both purposes are important ones for the individuals involved and the organizations they administer. Yet just as obvious is the difficulty the superintendent has in developing a helping relationship with that principal given the inherent threatening nature of the evaluation. Even a "good" evaluation carries an implicit threat, as Carl Rogers has noted,[5]

> .....In almost every phase of our lives - at home, at school, at work - we find ourselves under the reward and punishments of external judgments. "That's good"; "that's naughty." "That's worth an A"; "that's a failure." "That's good counseling"; "that's poor counseling." Such judgements are a part of our lives from infancy to old age. I believe they have a certain social usefulness to institutions and organizations such as schools and professions. Like everyone else I find myself all too often making such evaluations. But, in my experience, they do not make for personal growth and hence I do not believe that they are part of a helping relationship. Curiously enough a positive evaluation is as threatening in the long run as a negative one, since to inform someone that he is good implies that you also have the right to tell him he is bad.

As Rogers suggests, even though evaluations are necessary and "exhibit a certain social usefulness", the superintendent and principal in the example above may well experience anxiety or forms of role conflict, e.g., the inherent conflicts in being a helper and an evaluator simultaneously in the superintendent's case or being helped and evaluated in the case of the principal. Role expectations in a helping relationship seem to conflict with those of an evaluative relationship, sometimes producing feelings of uneasiness. Principals and teachers have often felt this same tension in conferences focusing on the evaluation of instruction. In order for evaluation and

development to become more compatible, it is necessary for them to be employed in ways to reduce the tensions associated with this apparent incongruence or role conflict.

In addition to the anxiety that often emerges from an inability to the integrate professional development and evaluation processes, tensions often arise because the process is perceived as threatening to the norms and sentiments of professional autonomy in educational organizations. For instance, teacher feelings or sentiments against external interference in their classroom operations[6] and administrators' beliefs and sentiments that support a high degree of autonomy in the operation of their institutions[7] have often been sources of tension in supervisory relationships. Certainly there are many functional aspects with respect to autonomy feelings among teachers or administrators. Teacher autonomy norms allow teachers to make daily decisions that would be difficult to orchestrate from a central office; the same is probably true of decisions administrators make in their own organizations and the difficulty a central office would have in making or coordinating these decisions. Further, how long could administrators maintain the confidence of their teaching faculties, if each time a decision needed to be made, the administrators were seen running off to call the central office for advice? Given these sentiments about autonomy, it is not difficult to understand possible tensions administrators feel when their immediate supervisor offers "advice" in a supervisory conference.

These anxieties, associated with the organization of professional development and evaluation in educational organizations, often lead to resistance which makes both data collection and analysis difficult. This resistance takes many forms: cynicism, antagonism, defensiveness or apathy

about professional development and evaluation. For instance, an educational leader's subordinates or colleagues may be reluctant to give information during appraisal which might "hurt" that leader's evaluation. This would often be so even though these same sources felt such information might be helpful for that leader's professional growth. Those giving information, even those responsible for giving feedback about professional performance of that leader, may also be inhibited by feelings that such input might be perceived as an intrusion with respect to norms about professional autonomy. Often data collection mechanisms or instruments avoid asking substantive questions as a way of avoiding conflict. In the end, the data are often compromised and neither professional development nor evaluation are successfully accomplished. To avoid these problems, there is a need to reorganize professional appraisal so that accurate data are available to support professional and organizational growth. In a word, that is what this text is all about.

Systemic Relationships

Appraisal can be viewed as a system relevant to the development and evaluation of professional performance. Indeed, the processes of appraisal in evaluation and appraisal in development can be viewed as two basic subsystems of an appraisal system, each systemically related to the other. For instance, in the human body, the cardio-respiratory system is composed of two major subsystems: the cardio-vascular system and the respiratory system. Each subsystem provides a different function. The cardio-vascular system, the heart and blood vessels, moves blood containing oxygen to various parts of the body and carries away carbon dioxide and other waste products. The respiratory system, the lungs and airways, conducts an exchange consisting

primarily of oxygen going into the blood stream and carbon dioxide moving out of the blood stream. They are both systemically related and functionally differentiated subsystems within a larger system to maintain life. In fact, they are systemically related and functionally differentiated subsystems with respect to virtually all subsystems or systems of the human body. At least in part, this analogy may be useful in describing systemic appraisal relationships.

Not to carry the analogy too far, but for illustration, educational organizations might consider the possibility of organizing appraisal for evaluation and development as functionally differentiated but systemically related subsystems of a larger system of professional appraisal which exists to promote individual and organizational growth. Systemic appraisal patterns would be related by a common concern for improving the performance of leadership personnel in the organization. The more congruent the definition of leadership performance held by each subsystem, the more systemically related would be those subsystems. For example, an appraisal system might be set up to focus on certain leadership behaviors relevant to particular organizational roles, these behaviors then would be the focus of professional development and evaluation. In essence, the systemically related appraisal system would provide diagnostic data to help professionals develop the same leadership behaviors that evaluation would address. Or to put it another way, it would also provide evaluative data on those behaviors that leaders were working to improve. Further, since the system of appraisal directly addresses goal-oriented leadership, this system and its subsystems could be viewed as vitally relevant to virtually all other organizational systems and subsystems.

As such, systemic appraisal in educational administration is not a new set of assessment instruments (instruments are probably already available) or a new set of supervisory skills (adequate supervision skills are frequently and effectively described in contemporary literature). Systemic appraisal is a way of organizing appraisal in the evaluation and development of educational leaders. Specifically, it features and emphasizes integrated appraisal of performance as a means for professional and organizational growth.

Systemic appraisal views the processes of evaluation and development as forms of rational thinking or decision making. In a developmental sense, appraisal is presented as a diagnostic, relatively nonjudgemental system by which educational administrators gather data about their professional performance so that they can make rational decisions about their own professional growth. In an evaluative sense, appraisal is presented as a data gathering system by which the organization can make rational and necessarily judgemental decisions about its educational leadership personnel.

A Comparison

As a means of further explaining the conceptual dimensions of systemic appraisal patterns, it might be helpful to compare and contrast its characteristics with more familiar, contemporary appraisal practice. Contemporary appraisal, here called "serendipitous" appraisal, typically emphasizes the evaluation process in the hopes that the professional's involvement in the process will result in serendipitous changes in performance. The relationship between evaluation and development is often an appraisal objective rather than an operational appraisal plan. Contemporary educational organizations run workshops, in-service training or cooperative relationships with colleges

and universities. These development activities rarely are in direct response or relationship to the appraisal process. Resulting positive changes in professional performance are more often than not credited to serendipity rather than any interactive or systemic relationship between evaluation and development. Certainly there are innovative attempts at appraisal that may be more systematic, less serendipitous in attaining goals, but for purposes of comparison, contemporary routine is emphasized at the expense of the innovative exceptions to the rule.

Table 1 attempts to compare and contrast the characteristics of serendipitous and systemic appraisal relationships between evaluation and development. Serendipitous relationships are characterized by: a relatively loosely structured relationship between evaluation and development; a single data base for both processes (recall the evaluation conference and suggestions for improvement); the predominance of the evaluation process; infrequent and sporatic ordering of evaluation and development events; definition of professional performance that is often vague and coercively imposed; appraisal processes that are almost always initiated by the organization; and because evaluation and development are rarely definitive, organizational incentives and rewards are generally applied. Systemic appraisal relationships on the other hand feature: evaluation and development in a deliberately planned relationship; a clear and consensual definition of professional performance; each process uses its own data base; the individual or the organization can initiate the processes which are continuously implemented; individual ownership of developmental or diagnostic data; equal weighting of both processes; and since evaluation and development are definitive, incentives are particularistically applied (outstanding performance is verified and rewarded and incentives are applied to improve performance deficiencies).

	Serendipitous Relationships	Systemic Relationships
Organization	Development and evaluation are related through serendipity rather than planning.	Development and evaluation are related through deliberate planning.
Performance Definition	Vague and often coercively imposed.	Clear definition built on consensus.
Data Base	Development and evaluation use the same data base.	Development and evaluation use separate and mutually exclusive data bases.
Initiative	Supervisor or organization initiates the processes.	Individual and organization can initiate the processes.
Data Ownership	The supervisor or organization owns the data.	The professional owns developmental data; the organization owns evaluative data.
Subsystem Equality	Evaluation is the predominant process; development is second.	The processes of evaluation and development are of equal importance.
Continuity	Evaluation and development are often sporatic and infrequent events.	Evaluation and development are frequently and continuously implemented.
Motivation	Organizational rewards and incentives are often generally applied and only indirectly related to the development process and performance.	Organizational incentives are particularistically applied and directly related to the development process and performance.

Table 1. A Comparison of the Characteristics of Serendipitous and Systemic Appraisal Relationships.

Comparisons can only go so far in describing systems, it is now appropriate in our discussion to focus on the operational dimensions of systemic appraisal. Three major areas of description: 1) the developmental appraisal subsystem, 2) the evaluative appraisal subsystem, and 3) the point of systemic linkage are the foci of discussion. As an example of how this system of appraisal could be applied to a particular group of educational leaders, the author focuses on possible systemic appraisal for school administrators or principals. Again, this specific example by no means is meant to suggest that systemic appraisal is relevant exclusively to principals, but only that principals are one of many groups of educational leaders, ranging from teachers to university presidents, that might wish to implement this model in some way to improve themselves or their organizations.

Development

Figure 1 presents a diagram of the appraisal subsystem for professional development. In this presentation, it is specified to the appraisal of school principals. Principals initiate and conduct their own appraisal and development activities. Further, the principals own the data they generate and are primarily responsible for using the data for developing and implementing professional growth activities. Data from the developmental subsystem can never be used as part of the evaluation process.[8]

The appraisal subsystem for development emphasizes the principals' complete ownership and control of the process as a means of reducing threat to professional security and autonomy. Since data collected for professional development can never be used in the evaluation process, participants in appraisal have little reason to distort or inflate assessments of performance.

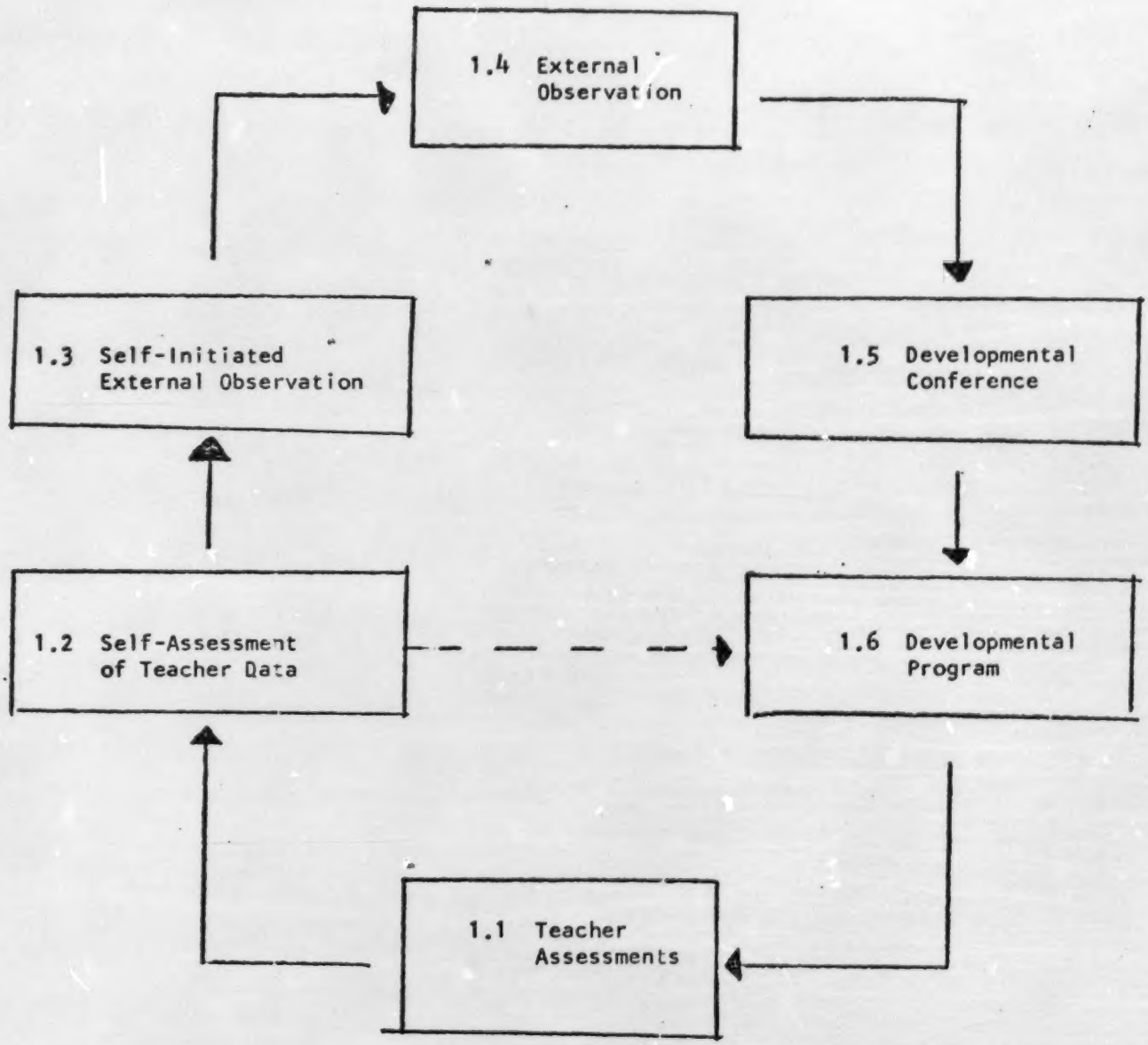

Figure 1. Professional Development Subsystem for School Administrators.

The data is simply used by the principal to improve both strengths and weaknesses based on diagnosis. To give inaccurate data or for the principal to cause data inaccuracies would only open up the possibility of a somewhat traumatic evaluation subsystem experience.

1.1 Teacher Assessments

The professional development appraisal subsystem or cycle begins with a diagnostic scanning of the principal's performance based on subordinate or teacher ratings. The assessments would be conducted at regular intervals, quarterly, biannually or annually by the principals themselves or by a central data gathering source in the school district with the trust of all involved. For instance, principals could select a data collector in their schools to distribute and collect the assessment instruments. The data collector would be someone that teachers know will maintain their anonymity and that principals know will keep the assessments confidential. Once all instruments are completed, the data collector returns them to the principal unscored. The principal scores the instruments and builds the diagnostic profile. In this way, only the principal knows the findings and only the principal owns the data.

The data could also be collected by a central agency, e.g., principals' association, a leadership academy, a special central office unit, a regional services center like BOCES in New York. While this organizational pattern has some advantages such as a centralized data base and reduced paperwork for principals, the principals would have to be assured that the data on specific principals would never be shared with their superiors or the public at large. Whether teacher data is collected in a centralized fashion or by principals

themselves, the point is that such data collection can be organized to provide accurate, helpful and confidential diagnostic profiles for professional development.

Yes, the gathering of teacher data may be initially threatening, perhaps even controversial. Many principals would argue that teachers are not an accurate source of information about administrative performance. Some would argue that the administration of organizational sanctions inevitably causes certain hostility that might negatively bias these data. In fact, it may well be that those principals that are just in the administration of sanctions, sooner or later will disappoint or anger everyone in their organization rather than a select group. While there is evidence to suggest that teacher assessments of the principals' performance do relate significantly to school climate and certain school outcomes,[9] many principals and supervisors of principals may always doubt the legitimacy of these assessments. For this reason, the teacher data is considered tentative. Anything suggested by teacher assessments of their principal is always subordinate to direct observations of the principal's performance by selected observers with administrative or leadership expertise.

1.2 Self-Assessment of Teacher Data

The diagnostic profile produced by teacher ratings of the principal's performance would include comparison of various dimensions of performance or comparison of individual performance with the average performance of all principals in the district.[10] After completing a personal profile, each principal could send an anonymous copy of his or her profile scores to each principal in the district. Identities would be hidden, yet averages could be determined. If a central agency were collecting the data, the agency could

easily compute these averages and supply them as part of the entire profile package. From these comparisons, principals can make tentative decisions about areas of strength or weakness upon which they wish to focus appraisal and development activities. For instance, a principal noting weak ratings on instructional leadership skills as compared to other performance areas and average ratings of principals in the district, might wish to set up observation and subsequent development in this area.

1.3 Self-Initiated External Observation

In a recent study of the principalship, principals tended to interact informally with other principals they thought had specific expertise about the principalship as well as demonstrated ability to maintain the confidentiality of the interaction.[11] This set of findings may be relevant to the means by which a principal initiates external observation. Two issues appear to become important in such a step: "Who shall do the observation?" or "What are the necessary characteristics of the observers?" and "What observation tools will be employed by the observers?" While possible instruments for the entire systemic appraisal system will be discussed later in this text, it may be helpful to focus on the essential characteristics of observers as a way of describing the principals' responsibilities in initiating external observation. Toward this general purpose, let's consider some theoretical propositions:[12]

P_1: As an external observer's understanding of the contextual idiosyncracy of a leader's system increases, so does the leaders confidence in the accuracy of the observer's appraisals of the leaders performance.

P_2: The more an external observer's status characteristics (expertise) render him/her more capable of appraising the performance of the leader than the leader him/herself; the more confidence the leader is likely to have in the accuracy of the observer's appraisals.

In other words, in order for outside opinions about the principal's performance to have an impact on that principal's self perceptions and future behavior, the observers must be viewed by that principal as understanding his/her particular work context or role and be seen as being better qualified to appraise that principal's performance than the principal him/herself. For example, a principal who as a result of teacher assessments and personal perceptions feels the need to appraise his/her pupil personnel skills might try to organize a three-person observation team composed of the following types of individuals: (1) another principal in the system with a thorough understanding of what it is like to be a principal in that district and a reputation of exemplary performance with respect to pupil personnel services, (2) a university professor familiar with that particular type of work context and expert in pupil personnel services, and (3) the principal being observed.[13]

P_3: There is an inverse relationship between the number of prior expectations held by the external observer for the leader's performance and the amount of confidence the leader has in the accuracy of the observer's appraisals of the leader's performance.

The principal must select observers, as P_3 notes, that have relatively few expectations (positive or negative) for the principal's behavior. To select a person who knows and in the past has either positively or negatively evaluated that principal's performance, may present the principal with doubts about the objectivity and accuracy of the observers' appraisals. While it may be impossible to always recruit total strangers, the principal's avoidance of those individuals who have a reputation for an inability to control prejudgements about performance is probably a wise strategy.

P_4: As the leader's trust in the observer's ability to keep his/her appraisals of the leader's performance confidential increases, so does the leader's willingness to accept the observer as part of the appraisal process.

The characteristic of confidentiality in the relationship between principal and observer is crucial to maintaining the nonthreatening dimensions of this part of the appraisal system. As noted earlier, principals routinely interact with peers they trust to maintain confidentiality.[14] In organizing the observations, the principal should emphasize the need for confidentiality.

Further, in organizing the external observation process, the principal should necessarily advise observers about times, places, areas of concern to which they can give intense consideration. This could probably be accomplished over the phone or in an initial meeting of the observation team. If necessary, training in the use of the observation instruments can be arranged. Recall, throughout this process, the principal is the initiator and organizer of the external observations.

1.4 External Observation

The actual observation of the principal's performance or the products of performance constitute the next step in the developmental appraisal subsystem. The observers should observe independently of each other; observers and observations can be spread over a week, or a month. The important aspect of these observations is the need to access multiple sources of data, e.g., interviews with various groups, documents, student achievement and attendance records. In gaining access to this information, the observers might consider the following propositions:

P_5: As the leader's (and significant other's) trust in the observers' ability to keep his/her appraisals of the leader's performance confidential increases, so does the observers' access to information essential to accurate appraisal of the leader's performance.

P_6: As the observers' access to information increases, so does the quantity and quality of the data available in the appraisal of the leader's performance.

As noted previously, the ability of an observer to maintain confidentiality and the subsequent trust of the principal may be a crucial observer characteristic, crucial not only to initial selection but also for access to data. Not only is it important that this be true of the observer - principal relationship, but it must be true with respect to the relationship between observers and members of the principal's organization. Nothing will increase the intensity of social defense mechanisms among teachers and others in the school more than the presence of a stranger asking questions. If observers are to be effective, the principal must prepare his/her organizational members for the visitors, testify to the observers' trustworthiness, inform organization members of the visits, and assure organizational members that their cooperation is needed and appreciated.

1.5 Developmental Conference

Once observations are completed and the observers have summarized the data and their thoughts, a developmental conference should be scheduled. Each observer must be present and the meeting should be held in a place that provides for a lack of interruption. The purposes of the conference are to give feedback and make recommendations for personal and professional development of the principal based on this feedback. The interaction might best be explained in terms of the following propositions:

P_7: As the quality and quantity of the data used in the appraisal of the leader's performance increases, so does the leaders confidence in the observers' appraisals of the leader's performance.

P_8: The more the observers' appraisals of the leader tend to focus on status characteristics (expertise) which favorably differentiate the observers from the leader, the more confidence the leader is likely to have in the accuracy of the observers' appraisals of the leader's performances.

The conference involves the principal and two external observers both of whom are perceived by the principal as having more expertise than the principal him/herself with respect to the targeted areas of observation. The principal has the responsibility of determining the nature of the data and data sources the observers used to complete their appraisals, and to keep the observers comments focused on those topics about which the observers can give expert advise and feedback. If these conditions are met, the following predictions or propositions might be helpful in understanding the subsequent interaction:

P_9: In the conference, if both observers make a similar or congruent appraisal of the leader's performance, the leader will tend to agree with that appraisal and respond appropriately in terms of personal and professional development.

P_{10}: In the conference, if the observers make different appraisals of the leader's performance, the leader will tend to distribute agreement between the observers and value these appraisals less than appraisals that have unanimous agreement (particularly in planning for personal and professional development).

Triad social interaction and the tendency toward coalition development is well known in the social sciences;[15] interaction patterns in the developmental conference would probably be similar. Given two observers or interactors that the principal believes are better able to appraise his performance than the principal him/herself, on those occasions in which both observers agree on a particular appraisal, the principal is likely to agree also. Even in the face of disagreement between the observers, the principal's agreement with either would produce appraisal decision. Clearly, in terms of the latter, the disagreement of "experts" may well diminish the value and impact of that

appraisal for development programming. These tendencies produce definitive decisions as well as decisions about needs that allow for the development of priorities.

In the developmental conference, the principal should take the leadership role, asking each observer to share assessments of performance, item by item. Once each observer has shared his/her perception, the principal would share his personal assessment and compare and contrast it to those of the observers. The process of developing generalizations about needs should be inductive in nature. Attempts at generalization should be reserved until all specific appraisals have been discussed. Discussion should be noncritical, allowing participants to express their opinions openly and without threat of rebuttal. Critical analysis of the data should take place at the end of the meeting. At this time, a general diagnostic profile serves as the central product of the meeting and the basis for personal and professional development.

1.6 Developmental Program

The principal would apply diagnostic data generated by the external observers to the creation of a development plan. The plan might best be articulated in terms of performance objectives.[16] These objectives would state conditions or learning alternatives, expected outcomes and criteria for successful completion. The school district might consider the development of an access capability to identify various relevant learning alternatives. This catalog or resource center might include a listing of alternatives, critiques by previous users, and programmed learning systems. Ideally, if a central data bank was available, cumulative data from diagnostic profiles could be used to evaluate resources listed or held in the center.

Over a specific time period, the principal would implement this development plan by taking courses, reading, counseling with others or whatever activities the plan specifies. The broken line in figure 1 moving from 1.2 to 1.6 notes the possibility that teacher assessment data might be available during the duration of the development plan. If this is the case, the principal might simply use these data as a partial feedback mechanism and modify or continue the plan accordingly. The principal always has the option of moving to external observation (1.2 to 1.3) or developmental programming after teacher assessment (1.2 to 1.6).

Evaluation

Since there is almost always an inherent threat involved in evaluation, the propositions (P_1-P_{10}) relevant to development might not initially appear applicable to the evaluation process. Certainly, during evaluation, the evaluator would be subject to a degree of defensiveness and data access would be limited to a greater extent than it would be as part of the development process. Often, the evaluator would hold prior expectations for a principal's behavior.

Keeping these considerations in mind, it would probably be a mistake to abandon the propositions just because they may be harder to apply in evaluation. Those doing evaluation appraisal must possess the same kinds of characteristics noted in the propositions presented above, i.e., contextual knowledge, specific expertise, a reduced number of prior expectations and the trust of those being observed. Those doing evaluations must work to establish this credibility among those being evaluated. With systemic appraisal, organizations that place people with questionable credibility in evaluator positions can no longer afford such luxury. Hopefully, systemic appraisal will provide an organizational

mechanism to hinder the "Peter Principle" circumstance that sometimes allows the least able to move to evaluation or supervisory positions.

With systemic appraisal patterns, development concentrates on those same dimensions of professional performance that evaluation addresses. Since the principal's immediate superior or supervisor knows that appraisal for development is being employed in the development subsystem, evaluation is the central issue. The supervisor makes "no bones about it," the supervisor evaluates performance and distributes rewards and incentives accordingly. There is no need to experience the development - evaluation role conflict.

If the principal chooses to use the evaluation as a cross-check of the development appraisal results he/she owns, fine, but the purpose of the evaluation subsystem remains solely in the realm of producing data by which the organization can make personnel decisions. While this may seem rather harsh and final, participants would soon note the absence of development-evaluation tension. Further, the system would also be systemically related to a grievance subsystem to resolve possible disputes between supervisor and principal through due process.

Figure 2 presents a diagram of the evaluation subsystem. It mirrors the development subsystem configuration and uses the same teacher assessment data employed in the development subsystem. In essence, 1.1 in the development cycle is in part 2.1 in the evaluation cycle. This is the point of systemic linkage between the two subsystems which will be discussed later.

2.1 Teacher Assessments

Instead of receiving specific teacher assessment data on each principal or a copy of each principal's diagnostic profile, the principals' immediate supervisor receives a district-wide profile based on principal "averages."

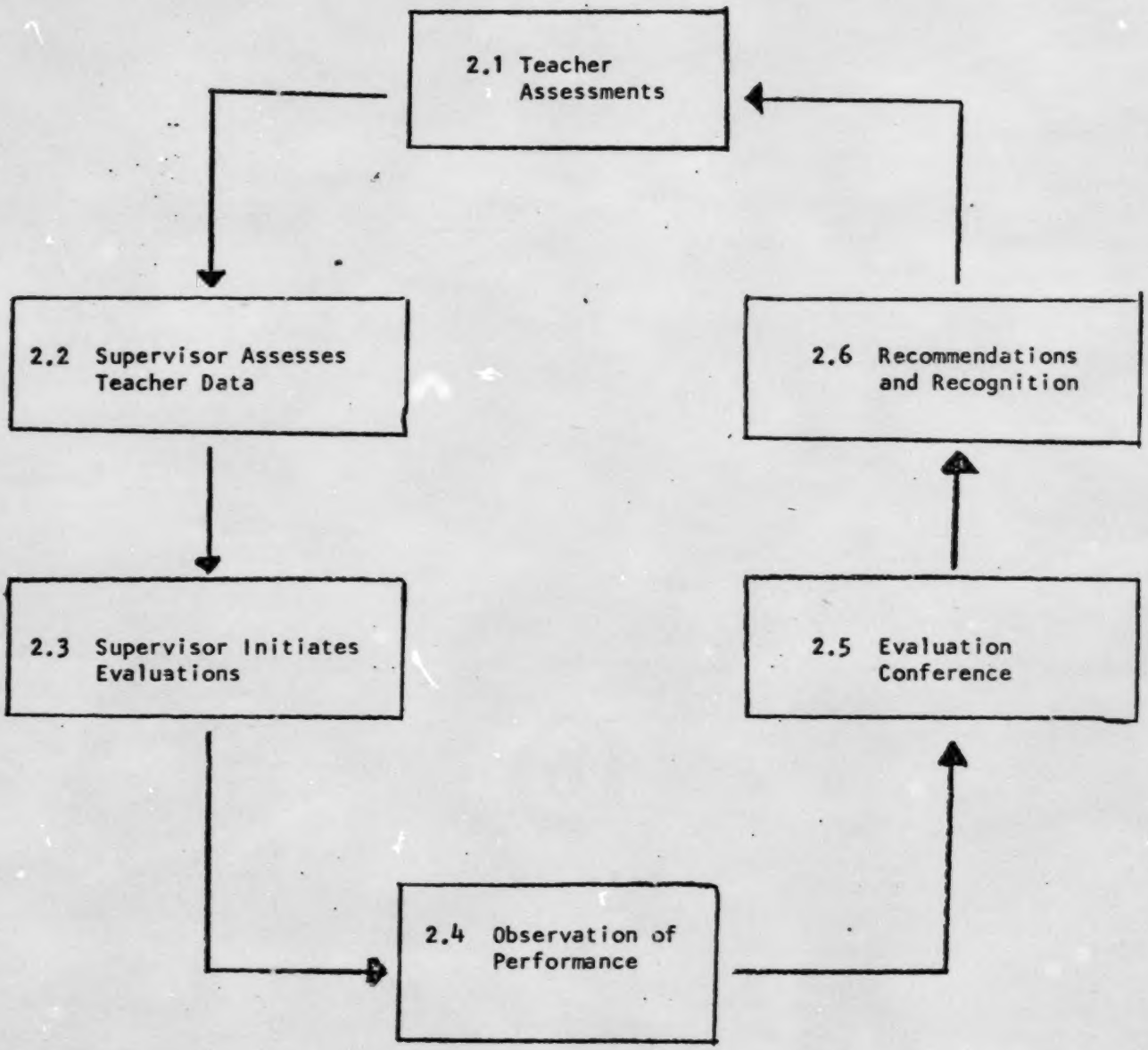

Figure 2. Professional Evaluation Subsystem for School Administrators.

The diagnostic profile would tell the supervisor about teachers' perceptions of the "average" principal in the district. To compute such a profile, principals would have to anonymously send the supervisor the numerical information used in computing scores for the individual profiles (an averaging of averages would not be sufficient). Ideally, this district-wide diagnostic profile would include various measures of central tendency and comparisons with assessments done at other times. If a central data bank existed in the system, the paperwork for the supervisor could be decreased. The purpose of this diagnostic profile is to graphically depict district-wide trends that need to be addressed and emphasized in evaluation.

2.2 Supervisor Assesses Teacher Data

Suppose the immediate supervisor notes on the district-wide profile that teachers in the district are relatively dissatisfied with supervisory relationships with respect to school discipline. This variable has the lowest average teacher rating as compared to other variables in the profile. Further, this variable rating appears to be declining over the past five years. On the other hand, instructional leadership skills received relatively positive ratings and these same kinds of ratings have been apparent over the past five years. Based on such information, the immediate supervisor might develop an evaluation strategy to determine whether or not these results could be verified. These areas of interest might be ripe for supervisor use of organizational rewards for excellent performance and incentives to improve questionable performance.

2.3 Supervisor Initiates Evaluations

Once the supervisor develops an evaluation strategy, the supervisor should arrange observations of the principals. These observations would employ

the same external observation instruments used in the developmental cycle. The most preferable type of external observation instrumentation would probably focus on products or formal physical evidence of performance.[17] Since the people in the organization may instinctively attempt to defend their principal in the face of a superior's evaluation, data based on testimony may not be as valuable as data based on products. It is this author's experience that teachers will tend to defend a principal against any outside evaluation - even if they personally dislike that principal or feel that the principal needs help. This tendency appears to be almost instinctive in nature, simply a cultural reaction to outsiders. Although, the principal should let teachers and others know of the supervisor's visit in advance and ask their cooperation in the evaluation process, one cannot assume such cooperation would necessarily be forthcoming.

2.4 Observation of Performance

Because of the resistance predicted above and in propositions 4 and 5, let's take a closer look at product evaluation emphasis. For instance, examination of a principal's documented plan for teacher supervision, the principal's written evaluations of teacher performance, the documentation of inservice efforts planned for teachers may be better measures of the principal's supervisory skills than teacher or principal testimony. Such instrumentation is available and will be discussed later.

The evaluation process would involve the principal being observed, the principal's immediate supervisor and a third party whose selection and presence would be agreeable to both the principal and the supervisor. This third party would necessarily possess relevant status characteristics, but could not be someone involved in the principal's development appraisal activities. If the supervisor wished to emphasize instructional skills,

people with a high degree of expertise in this area might be nominated as possible third-party observers. While this writer can easily think of a rationale for and against the use of a third-party, for field-test purposes, it might be wortn experimenting with this form rather than the more traditional one on one, supervisor-principal interacti on. The advantage of the third-party presence might be the reduction of problems inherent in an appraisal based on insider perceptions only.

Since the emphasis is on products, it might be helpful for the supervisor to advise each principal in advance of the products needed. This would not necessarily result in principals scrambling around at the last minute to develop products. Recall, these same products are part of the consensual definition of performance that the entire appraisal system addresses. It is quite likely that these products have been part of the principals' development programs in the past and are in place.

2.5 Evaluation Conference

While the introduction of this work seemed to malign evaluation conferences and much of the description to follow will seem similar to those in serendipitous patterns, there is a difference. This conference is based on data about specific performance, performance which is continually in the process of development. The conference is not a lone act of supervision, but part of a system which relates development to evaluation. In effect, principals are evaluated on those things they have been working to improve. Reduced to a degree is the role conflict of traditional evaluation conferences.

The supervisor should ask the principal to complete a self evaluation on the same product evaluation instrument employed in observation. While the

principal cannot share the developmental appraisal data he/she owns, this self-evaluation might be an indirect representation of such a data base. The interaction between supervisor and principal should focus on a comparison of the principal's self-evaluation with the supervisor's and the third-party's evaluations, item by item. This process is inductive and generalizations about performance or improved performance should be reserved until all the data is on the table. These generalizations are made solely by the supervisor based on an evaluation profile completed as a result of the conference. Recommendations about subsequent rewards and/or incentives for performance would be made later in the absence of the third party.

2.6 Recommendations and Recognition

One of the main features of systemic appraisal is its ability to distribute organizational rewards and incentives particularistically. In some school districts, regardless of evaluation results, everyone receives the same privileges, pay increase and recognition. With systemic appraisal, the supervisor now has a data base by which excellent performance can be rewarded and less effective performance can be identified. What rewards? These are limited only by the organization's resources and its leaders' creativity, e.g., special operating autonomy, funds for innovative programs, recognition or promotion. Necessarily, when all else has failed, the supervisor may need to make a personnel decision to dismiss an individual. However, in doing so, the supervisor and the individual knew in advance what was expected, specified development activities were continuously provided and time to correct problems was made available. As the last arrow in figure 2 shows, the cycle begins again with new teacher assessments (2.1).

Systemic Linkage

As noted above, the development and evaluation subsystems are systemically related through teacher assessment of the principals' performance. The linkage is more than this operational procedure. Required for such linkage is (1) a consensual definition of effective performance for principals and (2) instrumentation to operationally measure that definition. The two are necessarily interrelated.

Developing Consensus About Performance

As a precondition to systemic appraisal, system participants must develop consensus about a definition of effective practice. The definition necessarily would focus on performance descriptors that could be operationalized in terms of subordinate or teacher perceptions and products of performance. The definition should focus on those things principals do in operating an effective school. Operational definitions like, "Works with teachers to identify student needs" are preferable to statements like "Shows enthusiasm for the job." With the former, teacher ratings and a documented needs assessment relevant to students could be employed to substantiate this performance. In the latter, teacher rating might be relevant, but it might be difficult to find a product for "enthusiasm."

There are probably several ways to develop consensus, the delphi technique, the nominal group process and other organizational development processes would seem to be applicable to the task.[18] Everyone involved in appraisal, principals and their immediate supervisors, should be involved in the selection or development of the consensual definition of effective practice.[19] There are three ways to proceed. First, the participants could develop consensus on a general definition of performance such as the one presented in figure 3. Note the relative simplicity and generality of this sample definition. From this

Principals in this district must effectively and frequently perform the following as a means of operating an effective school:

1. Principals must provide leadership with respect to curriculum and instruction in their schools.

2. Principals must administer the staff personnel program in their schools.

3. Principals must administer the pupil personnel program in their schools.

4. Principals must implement system-wide policies and procedures in their schools.

5. Principals must administer the fiscal management of their schools.

Figure 3. Sample Consensus Definition

general definition, a committee of participants and/or consultants would then need to develop more specific, behavior descriptors of what a principal would do to accomplish those things noted in the consensual definition. So that these could be rated and scored in the development of diagnostic profiles, these behavioral descriptors would be linked with scoring scales, i.e., rating each performance on a 1 - 5 continuum, the higher the score, the better the performance. Because of the number of developmental tasks, this is probably the most complex and difficult means to an operational definition of effective performance.

The second possibility, would simply be to tentatively adopt both a conceptual and operational definition that has been previously developed. These definitions might be already in place in the organization or school district adopting systemic appraisal. There may be an instrument already being used to assess the performance of school principals. If it employs behavioral descriptors and has a sound theoretical base, why not use it or adapt it.

Often such definition and instrumentation is not available, particularly for school administrators. If this is the case, the participants should consider the possibility of tentatively adopting or adapting a definition and set of instruments that have already been developed elsewhere. As an example of such instrumentation, the Principal Performance Description Survey (PPDS) provides such a possibility.[20] These instruments measure the principal's performance in terms of teacher perceptions and external observer ratings of the products of performance. The items in the instruments are conceptually based on an extensive review of the literature, time-motion studies of practicing principals and verification by hundreds of practicing professionals.

These behaviors or items are known to be related significantly to teacher and student perceptions of school climate and to student achievement and attendance.[21] While these instruments may not be totally congruent with the nature of practice in the principalship from one district to another, this author would argue that there is more in common between these measures and differential practice in the principalship than there is different. Further, these instruments would provide at least a starting point for field-testing and adaption to a specific district. Items might need to be modified, some eliminated, some added, but at least the instrumentation is in hand and consensus can be built through practical trial and error rather than extended theoretical and philosophical discussions that often lead nowhere.

One other issue needs to be noted relative to consensus building. While consensus should be based on the input of all system participants, there probably needs to be an external cross-check on the consensus building process. Without such a cross-check or devil's advocate role in the process, those involved may overlook new ideas in favor of the status quo. An external consultant or an organizational member whose job it would be to play the devil's advocate could provide input to avoid "watered-down" or inadequate definition of effective performance.

Instrumentation

In place, the consensual definition of effective performance would provide a focus for the development and evaluation subsystems. Principals could employ appraisal and development activities to the same performance definition that evaluation would address. To do this, two operational definitions would be needed to provide measures for the consensual definition of performance. With principals, an instrument to measure teacher perceptions of their

principal's performance and an external observation tool to assess the products of principals' performance would need to be developed or identified.

As noted above, the Principals' Performance Description Survey, particularly the teacher form and the external observer form are examples of possible operational definitions. These instruments assess a principal's performance through teacher perceptions of the principal and external observer ratings of the products of the principal's performance. This battery of instruments addresses several functional areas of responsibility, i.e., curriculum and instruction, staff personnel, pupil personnel, system-wide policies and procedures and fiscal management (fiscal management is on the external observer form only). Also, as a way of showing how a consensus definition could be specified operationally, these PPDS measures could be possible examples of the operational measures of the definition noted in figure 3.

Items on the teacher form of the PPDS ask the respondents to rate both the "effectiveness" and "frequency" with which the principal performs certain tasks in the school. Each item is scored on a 1 - 5 continuum. The higher the score, the more effective or more frequent the performance is rated by respondents. Since high effectiveness and frequency ratings are known to be related to meaningful school outcome measures,[22] the higher the scores, the better the performance rating. All item scores under each functional area, e.g., curriculum and instruction, are totaled to produce a general measure for that area. Figure 4 presents a sample part of the teacher form of PPDS, specifically from the section on curriculum and instruction responsibilities.[23]

Recall, the teacher form would be used to present a diagnostic profile to each principal to initiate the development cycle. The profile would graphically compare teacher ratings of different functional areas of responsibility with

PERFORMANCE STATEMENT SHEET

RATING SCALE

Frequency With Which Principal Performs Task	Effectiveness With Which Principal Performs Task
1 2 3 4 5	1 2 3 4 5
NEVER ... VERY OFTEN	INEFFECTIVE ... VERY EFFECTIVE

CURRICULUM AND INSTRUCTION

1. Evaluates the instructional climate by observing in the classroom.
2. Works with teachers in formulating grading practices and procedures.
3. Encourages teachers to consider individual differences when evaluating student performance and progress.
4. Discusses changes in the educational program with teachers.
5. Encourages teachers to work together in planning and modifying the curriculum.
6. Informs teachers of general teaching practices and skills for which they are responsible.
7. Encourages teachers to try new and innovative teaching methods in helping the consistently failing student
8. Discusses problems of consistently failing students with teachers.
9. Discusses classroom goals and procedures with teachers.
10. Works with teachers in understanding and using results of the school testing program.
11. Plans a variety of instructional programs to meet individual learner needs.
12. Works with curriculum committees to establish educational goals of the school.
13. Works with faculty committees to review curriculum content and organization.
14. Works with teachers in establishing student performance standards.
15. Works with teachers in evaluating the classroom instructional climate.

Figure 4. Sample part of the PPDS teacher form.

one another or with average scores for all principals in the school district. Figure 5 presents an example of such a diagnostic profile. As part of the evaluation cycle, average teacher ratings for the entire school district on the various functional areas could be compared to one another or average teacher ratings from the most recent assessment can be compared to assessments done at other times. Figure 6 presents an example of a district-wide profile which the principal's immediate supervisor can use in planning an evaluation strategy. These profiles would serve as the systemic linkage between development and evaluation subsystems. In essence, principals will be working on the very same thing that their immediate supervisor will be observing during evaluation.

The external observer form of the PPDS focuses on observed evidence or products of a principal's performance. Each performance descriptor or item on this form must be substantiated in terms of a scale ranging from "informal" to "product (formal)" evidence of performance. Again, items measure a principal's performance in terms of the same five functional areas of responsibility that are addressed in the teacher form of the PPDS. A sample item and response scale that an external observer or evaluator might employ during appraisal is presented below:[24]

1. Keeps information about new research and methods in education on-hand for personal and staff use.

Source	NA	Informal	Input	Process	Product
_____	_____	1	2	3	4

Product Description _____

_____

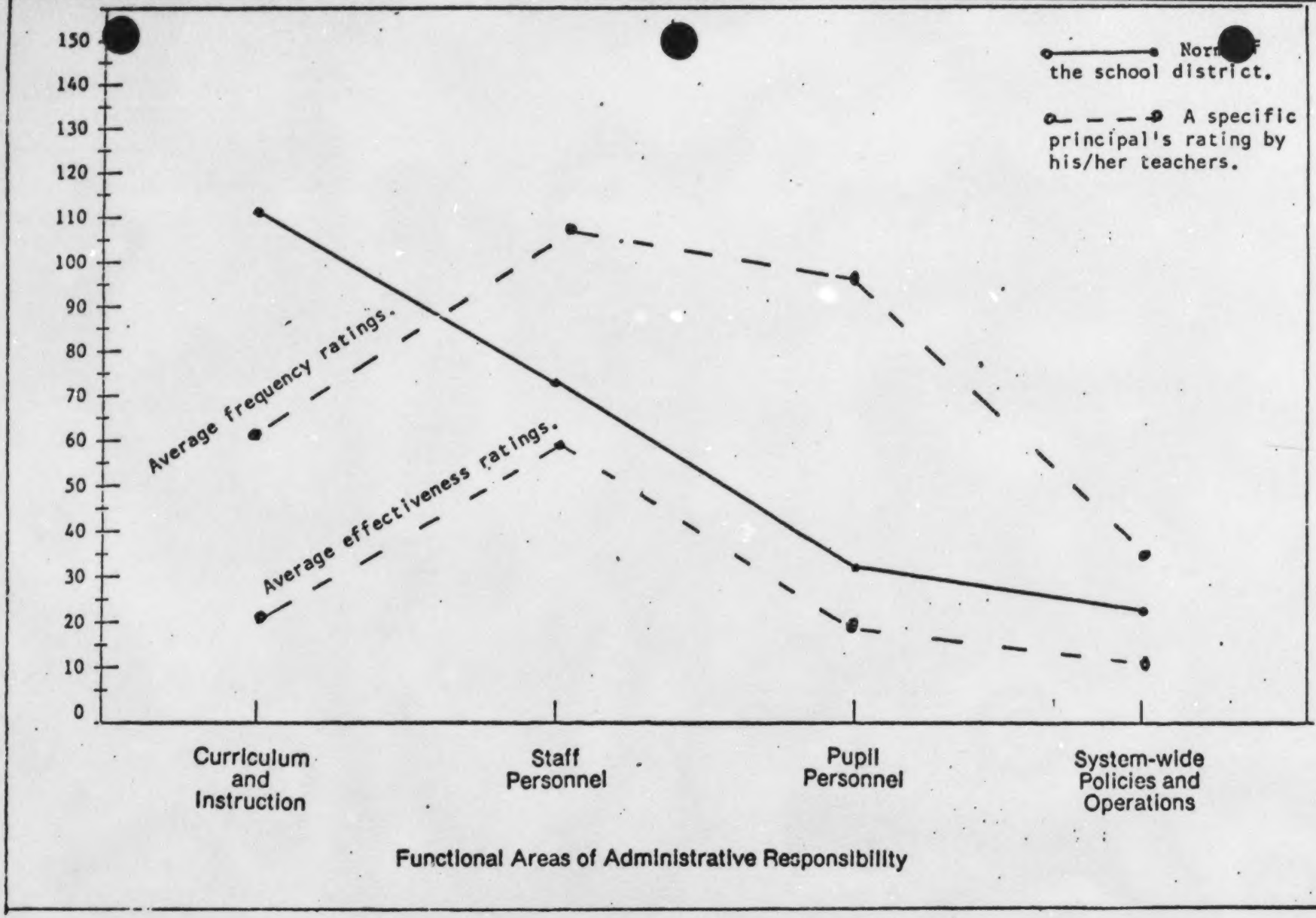

Figure 5. Sample diagnostic profile based teacher ratings.

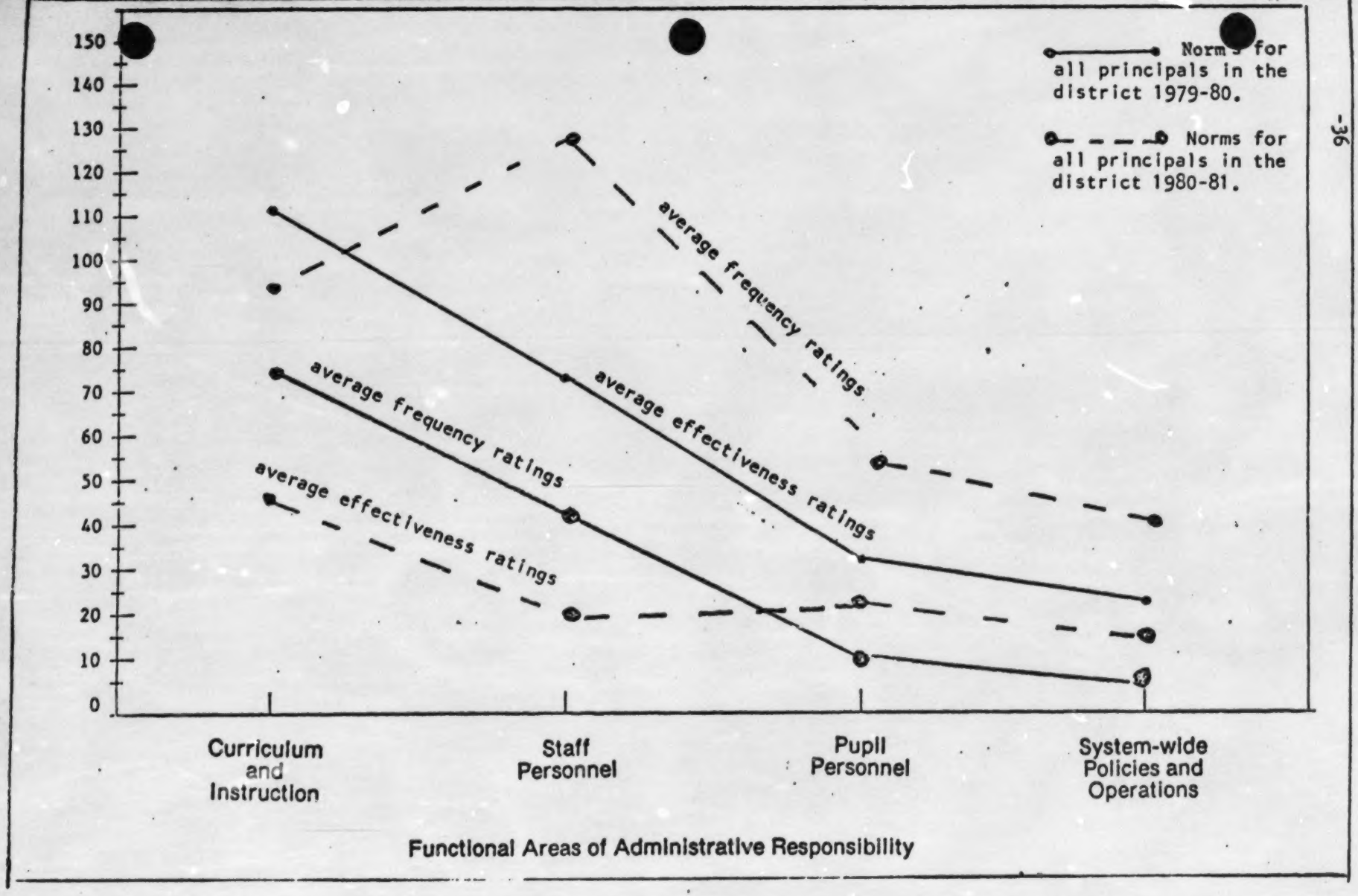

Figure 6. Sample evaluative profile based on teacher ratings.

Let's assume a principal is being interviewed by one of the external observers he/she has selected for the development cycle or the principal's immediate supervisor during the evaluation cycle (each using this external observer instrument). Once rapport is established and a general explanation of the procedure and how it is to be administered has taken place, the interview and observations may begin. While there are specific interview or observation questions on the form, the observer is not restricted to only questions on the form in trying to determine the meaning of the principal's response.

Scoring of each item or question involves the following procedure. If the respondent identifies another individual in the school who has responsibility for a given performance (e.g. secretary, an assistant, a department chairman, etc.), this person is to be noted in the space provided under "source." These persons are to be interviewed later. If the respondent understands the question but states that it is not applicable to principal's job description in that district, the observer checks the space under NA and proceeds to the next question or item.

Recall that in the process of producing a product score, the scale varies from 1 - INFORMAL to 4 - PRODUCT. If the respondent agreed that he performs a particular behavior, but could not provide the external observer with direct observable evidence, but claims it is done informally, he/she is given a score of 1 for the INFORMAL category. If he/she had on hand materials relevant to a particular performance, but had done nothing with them, e.g., personality profiles for students that had never been used in assessing student needs, the observer scores a 2 for INPUT. If, in the observer's judgment the respondent was in the process of carrying out a particular performance beyond

the input stage, e.g., asking teachers to critique personality profiles before use in policy making he/she is given a score of 3 for PROCESS. The highest score of 4 for PRODUCT is given for only directly observable evidence relevant to a specific criterion or procedure beyond the process stage.[25]

Under the product scale is an open-ended section entitled "Product Description." The observer is to use this section to note any and all informal or subjective information derived from the interview. For instance, if the respondent produces a list of advisory committee members and minutes of their last meeting, these products would be noted in this section as well as any informal comments that might be used in making qualification judgments. If, when asked about identification of qualified minority group candidates for teaching positions, the respondent reports that this is a "word of mouth" or informal process, a description of this informal process would go under Product Description. In this way the observer is able to get both verification of formal procedures in practice and specific information relative to informal processes. Figure 7 presents a sample part of an external observer form of the PPDS, specifically from the section on staff personnel responsibilities.[26]

This particular or external observer instrument would be completed by the principal him/herself prior to the observers' visits or at least prior to the developmental conference and the observers selected by the principal to participate in the developmental cycle. This instrument would be completed by the principal him/herself and evaluators during the evaluation cycle. While the PPDS does feature a self-rating instrument that can be used with both the teacher and external observer forms, this author's suggestion would be that in systemic appraisal the external observer instrument would provide the best vehicle for interaction and comparison during development or evaluation conferences. This would allow the teacher data to be kept separate from the observation process.

Staff Personnel

8. Matches employee qualifications and job descriptions in selecting new employees.

Source	NA	Informal	Input	Process	Product
_____	____	1	2	3	4

Product Description _____

_____

9. Maintains written job descriptions for employees.

Source	NA	Informal	Input	Process	Product
_____	____	1	2	3	4

Product Description _____

_____

10. Informs teachers of guidelines to be followed in reporting student disciplinary problems.

Source	NA	Informal	Input	Process	Product
_____	____	1	2	3	4

Product Description _____

_____

11. Evaluates staff participation in formulating school policies and procedures.

Source	NA	Informal	Input	Process	Product
_____	____	1	2	3	4

Product Description _____

_____

12. Maintains written policies concerning school rules and regulations for students and teachers.

Source	NA	Informal	Input	Process	Product
_____	____	1	2	3	4

Product Description _____

_____

Figure 7. Sample part of the PPDS external observer form.

Figure 8 presents a sample profile that might be the result of external observation in either the development or evaluation subsystem.[27]

The Entire System

The entire systemic appraisal system is presented in figure 9 as it applies to school principals. This graphic presentation illustrates the systemic relationship between the two functionally diferentiated appraisal subsystems of development and evaluation. As noted above, teacher assessments of the principal (1.1 in development; 2.1 in evaluation) serve as the systemic point of linkage between the two subsystems. Recall, the system applies to principals in this example, but it could be applied to other leadership roles.

Teaching Faculty

If we can view classroom teachers as instructional leaders of students, the systemic appraisal system could be employed as a professional support system for teachers at the school building level. Certainly, with many available measures of teaching effectiveness, a consensus definition of performance is possible. A systemic appraisal system for teachers is presented in figure 10. In this application, the department chairperson or the building principal would be the evaluator, student assessments of teaching are the systemic linkage point.

Deans of Colleges

Moving to a higher education context, the system might be applied to the academic leadership of a college or university. For instance, figure 11 presents such an application for college deans in a large university. If a consensus definition of effective performance could be identified and measured, professional faculty assessments would probably serve as the point

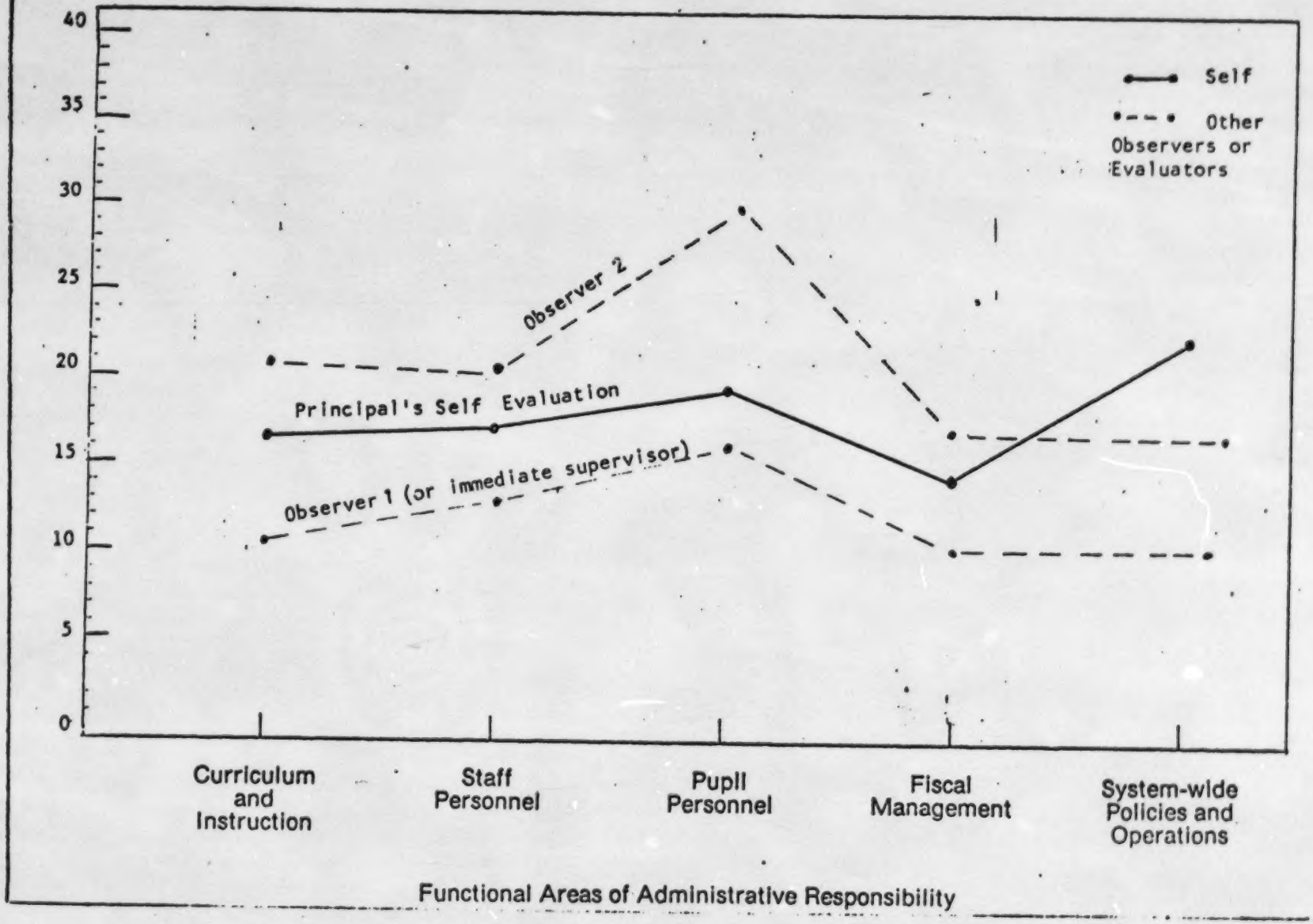

Figure 8. External observer profile from evaluation or development subsystems.

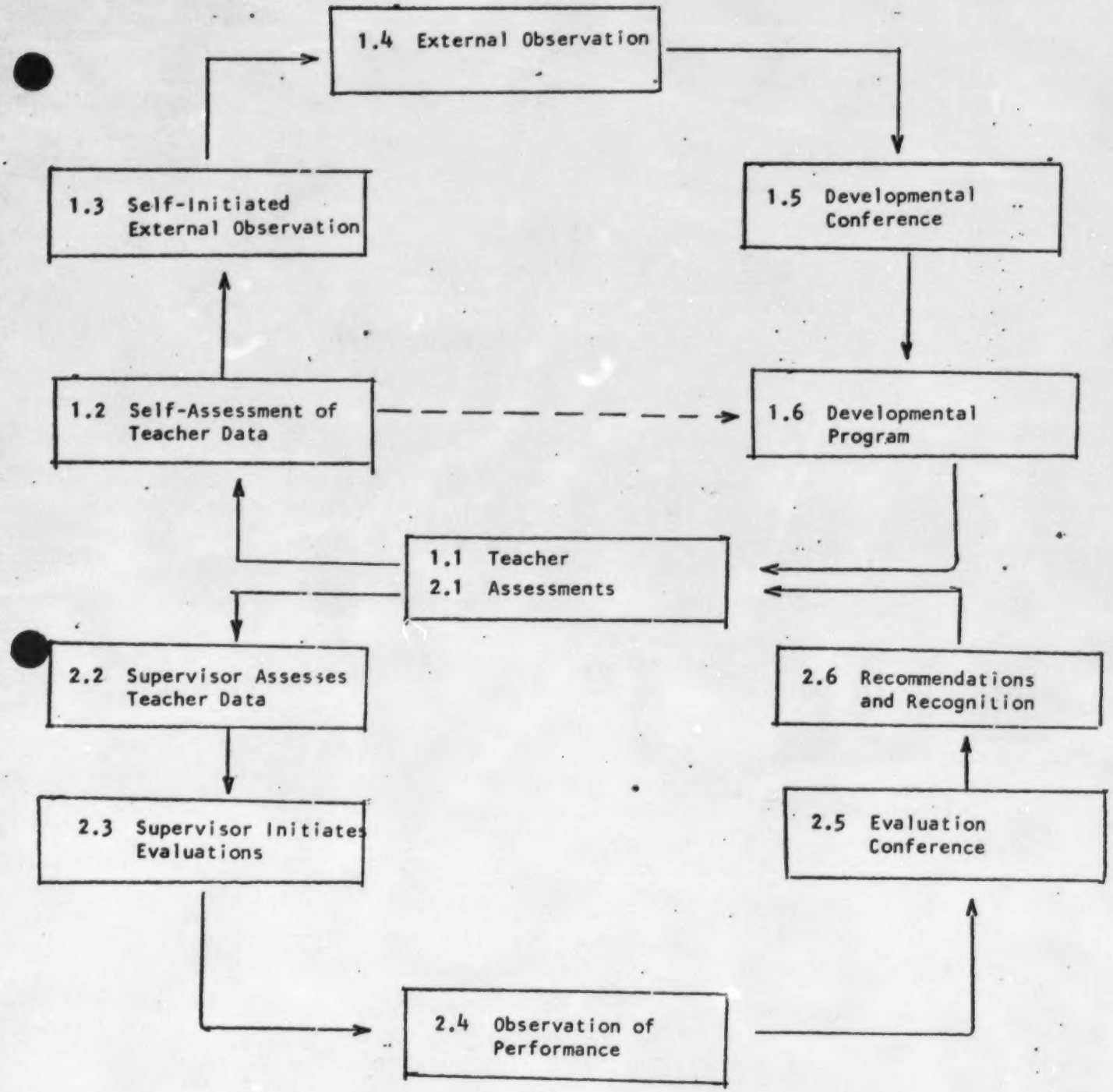

Figure 9. Systemic appraisal for school principals.

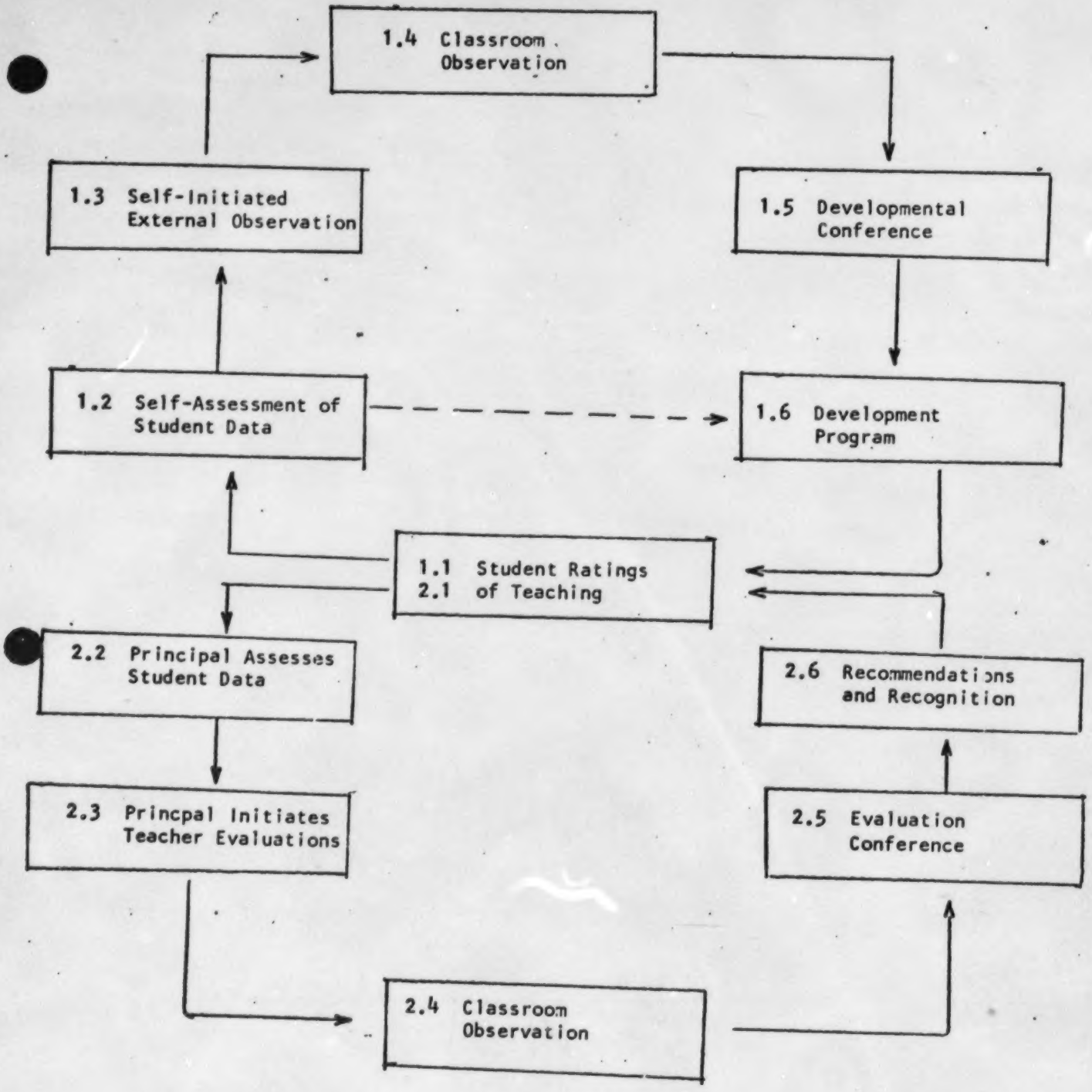

Figure 10. Systemic appraisal for teachers.

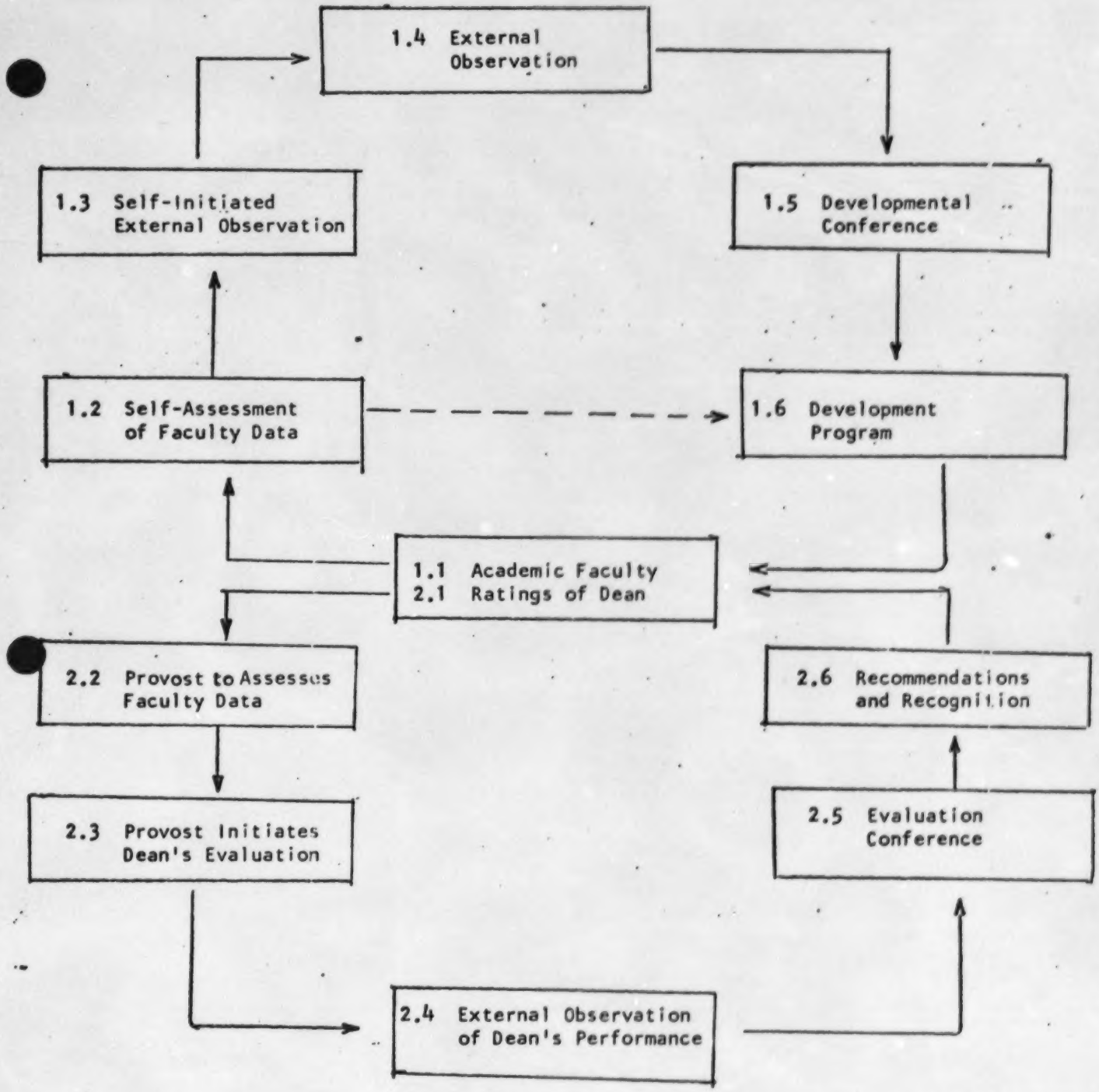
1.4 External Observation
1.3 Self-Initiated External Observation
1.5 Developmental Conference
1.2 Self-Assessment of Faculty Data
1.6 Development Program
1.1 Academic Faculty
2.1 Ratings of Dean
2.2 Provost to Assesses Faculty Data
2.6 Recommendations and Recognition
2.3 Provost Initiates Dean's Evaluation
2.5 Evaluation Conference
2.4 External Observation of Dean's Performance

Figure 11. Systemic appraisal for college deans.

of systemic linkage. The university officer such as the provost might serve in the evaluation role.

Division Heads in Government or Industry

The system might be applicable to leadership in government or industry; division heads might be examples. Suppose a large corporation employed a number of division heads to provide leadership with respect to various specialized corporate functions. Systemic appraisal might be employed using division employee ratings as the systemic link between development and evaluation. While functions of each division head might require specialized expertise, a consensus definition of generic leadership behaviors applicable across the board probably could be developed. The division heads' direct superior, perhaps a corporate vice president would serve in the role of evaluator. Figure 12 presents this adaptation.

Executives

The systemic appraisal approach is most easily applied to leadership positions in an organization that has several role incumbents. When there exists only one individual in a leadership role, subordinate assessment data would always identify the subject of assessment. No longer would the evaluator receive averages of subordinate ratings, but specific ratings for a single individual. This probably would increase the potential threat of such data as well as increase doubts and objections about the reliability and validity of the data. Executives like university president, school superintendents, or coorporation presidents may confront such a problem in adapting systemic appraisal to their positions.

Does this mean that systemic appraisal is applicable to every leadership

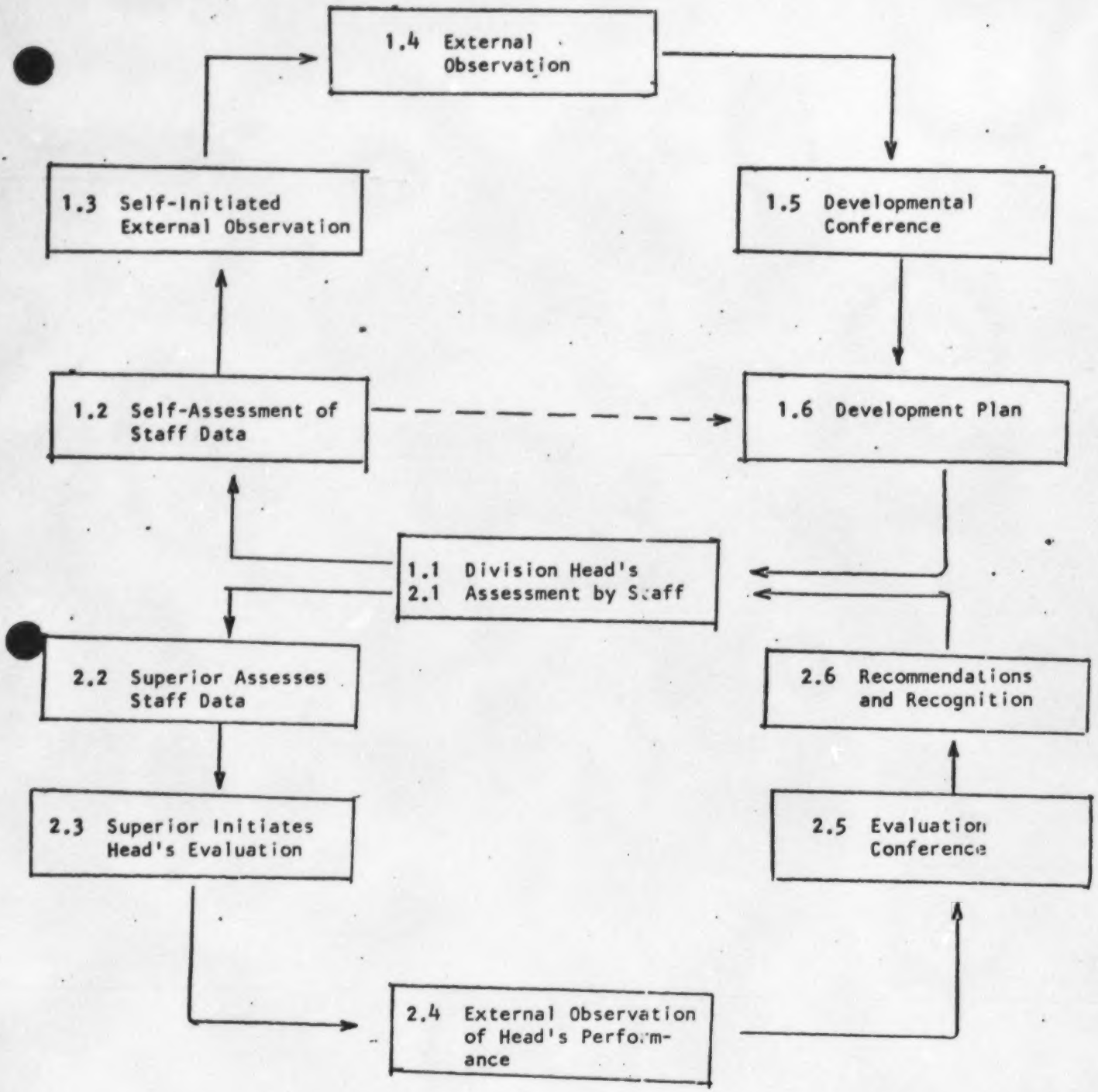
1.4 External Observation
1.3 Self-Initiated External Observation
1.5 Developmental Conference
1.2 Self-Assessment of Staff Data
1.6 Development Plan
1.1 Division Head's 2.1 Assessment by Staff
2.2 Superior Assesses Staff Data
2.6 Recommendations and Recognition
2.3 Superior Initiates Head's Evaluation
2.5 Evaluation Conference
2.4 External Observation of Head's Perfo:m-ance

Figure 12. Systemic appraisal for division heads in government or industry.

position but the highest in authority? Are only subordinate leaders accountable for their performance? Are there leadership positions so elite that they defy appraisal and professional growth possibilities? Perhaps, The Pope or the President of the United States might be such examples (although both are accountable for their behavior). This writer recognizes the existence of such exceptions, but rejects, for the most part the notion that systemic appraisal is not applicable to executives. Most are subordinate to some governing body, few are absolute or "god-like" authorities and most need to address professional growth as a way to meet the challenges of their positions.

Adaption of the systemic appraisal approach to executives might include use of subordinate assessments as part of the evaluators data base with those executives who are not threatened by such information and procedures. Here, special care would be taken to maintain the confidentiality of the data base. In the case of an executive who reports to a governing board, a subcommittee might handle evaluation. Where subordinate ratings could not be used in the evaluation subsystem, the systemic link would be the consensus definition of performance rather than consensus definition and subordinate ratings. In this case, the executive would still receive a diagnostic profile of subordinate ratings from the development subsystems, but no subordinate data would be available to the evaluator(s). External observation or product data would be the central source of data for evaluation.

Another approach to systemic appraisal for executives might be the development of executive consortiums in which the state or nation (rather than the organization) is the unit analysis. While this is probably a controversial suggestion, superintendents of schools and/or college presidents within a state would constitute a consortium for professional growth.[28] In order to successfully develop such a structure, a statewide consensus

about the definition of executive performance would be a challenging undertaking. Using the state as the unit of analysis and the goal of improving executive leadership within the state as its central purpose, systemic appraisal could be implemented with a state-wide data base. Each executive would receive and own a private diagnostic profile of subordinate ratings comparing and contrasting individual performance with state-wide averages. Evaluators, perhaps a team composed of local and state officials, would receive a profile of regional or state-wide subordinate ratings to help develop an evaluation strategy. External observation in development and evaluation would not only provide data on individual executives, but another on-going data source about executive leadership in the state.

Implementation

The history of innovation in education, particularly the implementation of innovation in educational organizations, has been characterized by less than successful efforts in many circumstances.[29] Certainly, the literature about possible pitfalls as well as strategies abound.[30] It is not the purpose of this work to discuss implementation strategies at length with respect to systemic appraisal applications in educational organizations. This author would recommend works like those by Havelock[31] or Bennis, Banne and Chin[32] to the reader to address this problem. Instead, this author will simply sketch out the basic functions or outcomes that might be essential in a successful implementation effort. These functions are presented in a flow chart in figure 13.[33]

A successful implementation effort would need to accomplish at least five basic functions or outcomes: (1.0) a decision by system participants

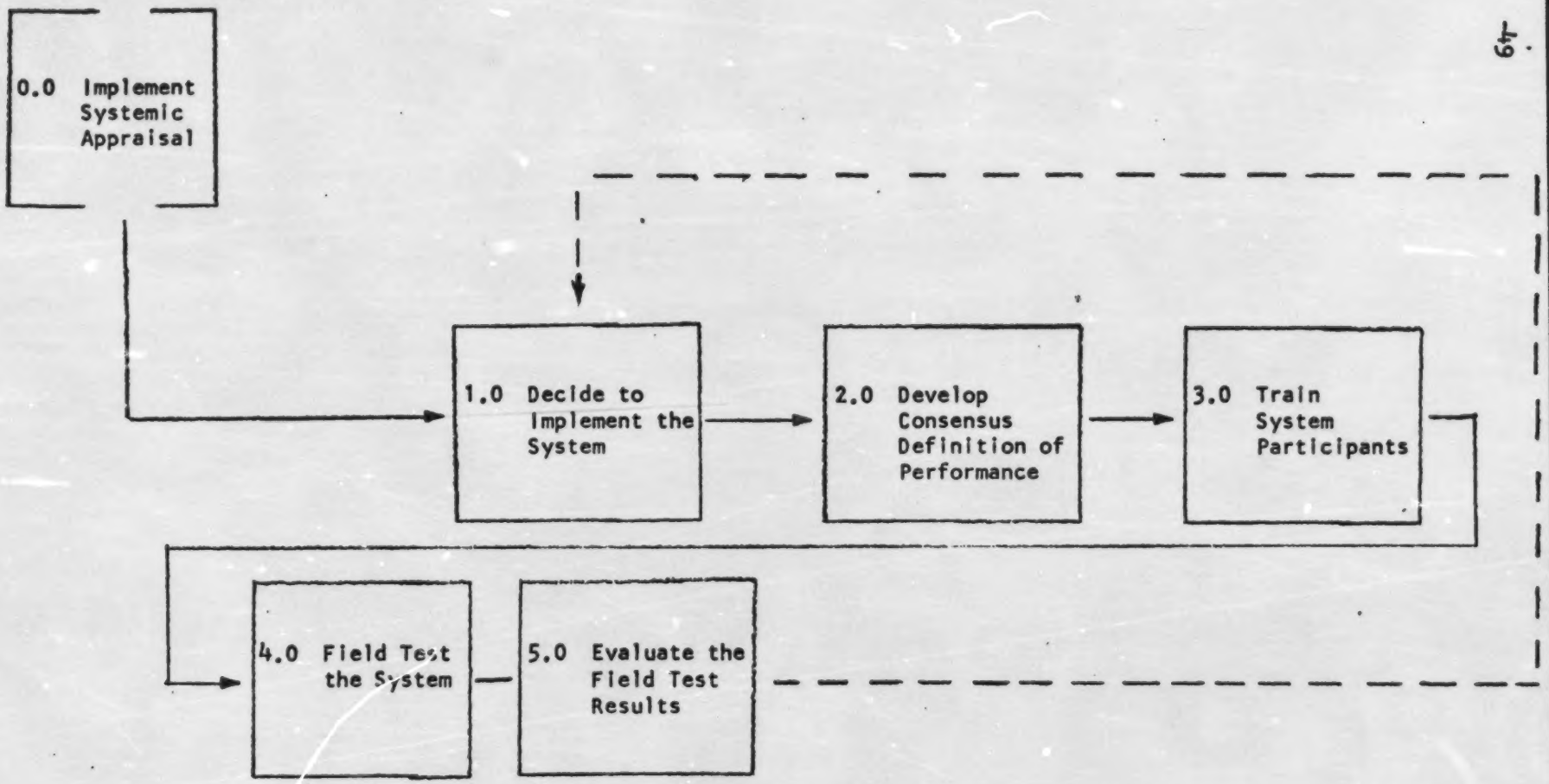
0.0 Implement Systemic Appraisal
1.0 Decide to Implement the System
2.0 Develop Consensus Definition of Performance
3.0 Train System Participants
4.0 Field Test the System
5.0 Evaluate the Field Test Results

Figure 13. Flow chart of the general functions associated with implementation of systemic appraisal.

to develop the appraisal approach in their organization; (2.0) a consensus definition of performance; (3.0) the training of system participants with respect to the knowledge and skills necessary to run the system; (4.0) an actual field test of the system; and (5.0) an evaluation of the field test results as a means to system improvement. These functions will be described in more detail in the following paragraphs.

(1.0) Decide to Implement the System

If one were to consider the implementation of this system in a particular organization, a study of both the informal and formal organization would be necessary to identify opinion leaders and organizational needs (1.1). When implementation processes exclude influential organizational members or fail to address the expressed needs of possible system participants, success would be doubtful. Once the opinion leaders are known, a committee or committees could be developed to consider the appraisal system as an alternative (1.2), particularly in terms of identified needs (1.3) and the consequences of systemic appraisal in addressing those needs (1.4). Given sufficient time to examine systemic appraisal as an alternative, the committee would ascertain the level of interest among possible participants and organizational leadership (1.5) as a means of making a decision about whether (1.6) or not (1.7) to try to implement. The broken line or feedback loop suggests that regardless of the decision, continued examination of needs and structure is needed. Unless the alternative has the wide-spread support of organizational leadership and membership, a decision to implement would be unwise. Figure 14 presents a flow chart of these functions.

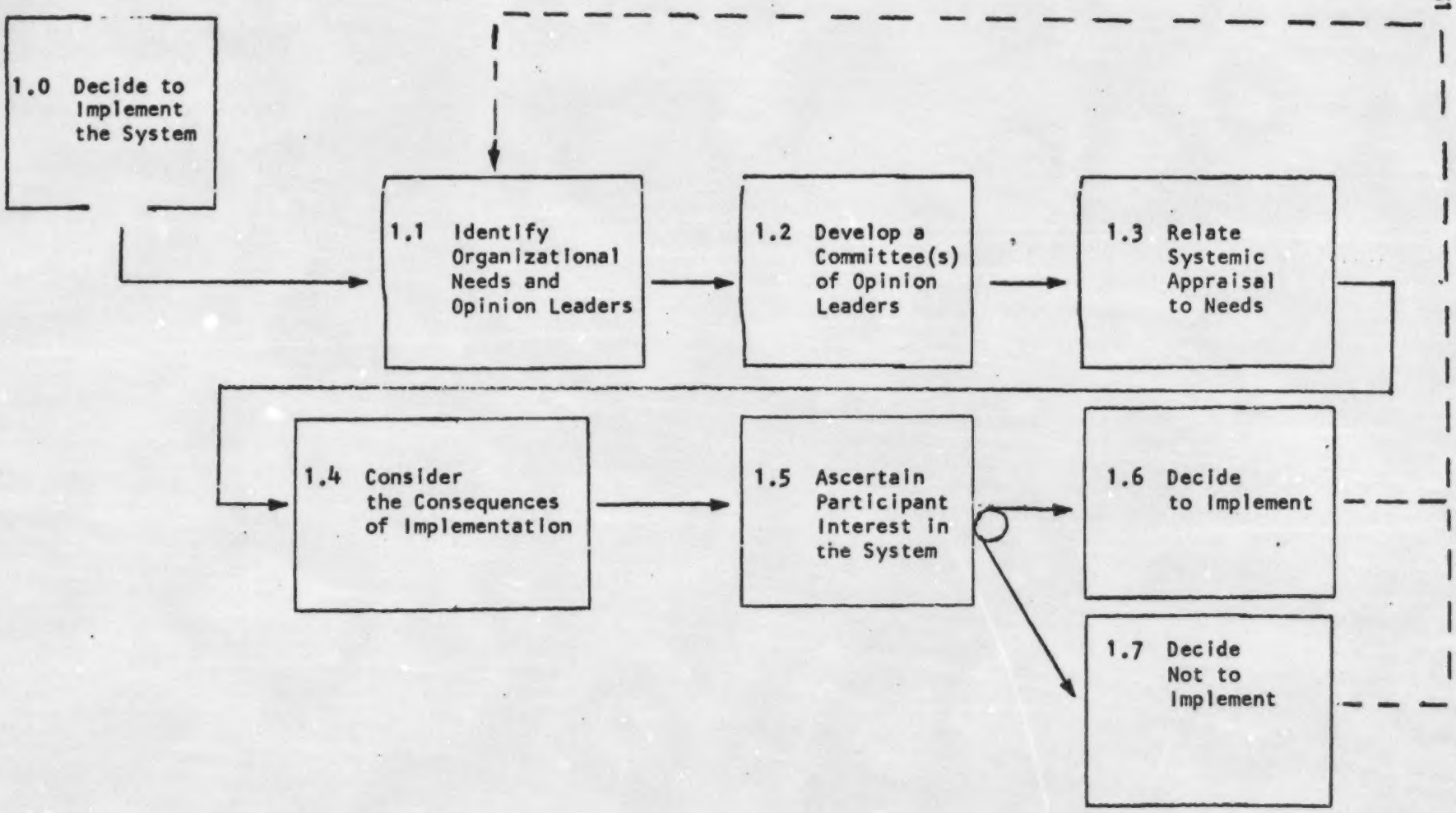
1.0 Decide to Implement the System
1.1 Identify Organizational Needs and Opinion Leaders
1.2 Develop a Committee(s) of Opinion Leaders
1.3 Relate Systemic Appraisal to Needs
1.4 Consider the Consequences of Implementation
1.5 Ascertain Participant Interest in the System
1.6 Decide to Implement
1.7 Decide Not to Implement

Figure 14. "Decide to Implement" flow chart.

(2.0) Develop a Consensus Definition

Figure 15 presents a flow chart of possible functions necessary in consensus building with respect to performance definition. As noted previously, nominal group processes or the delphi seem appropriate to this task. Using the committee or committees of organizational opinion leaders, consensus about a conceptual (2.1) and operational definition (2.2) of performance needs to be accomplished. Operational definition would necessarily include specification of data sources (2.21), identification or construction of appropriate instrumentation (2.22) and the organization of an operational information system to manage effective storage and flow of data (2.23). No definition should be "set in concrete", always considered tentative based on field testing and evaluation (2.3). As the broken line or feedback loop notes, evaluation leads to new consensus.

(3.0) Train System Participants

Every system, in order to survive, must be understood by its participants and taught to new participants. A training program would need to be set up, perhaps in the form of a manual or in-service education (3.1). This training program would explain the procedures of data gathering (3.11); identify a pool of external observers and necessary observer characteristics (3.12); and identify resources that can be employed in responding to diagnosed needs (3.13). The training of participants in analysis of data, building profiles and generalizing about needs would constitute another component of training (3.2). A component that would evaluate the effectiveness of these training procedures (3.3) could be employed as a way of improving the program (see feedback loop). Figure 16 describes the flow chart for these functions.

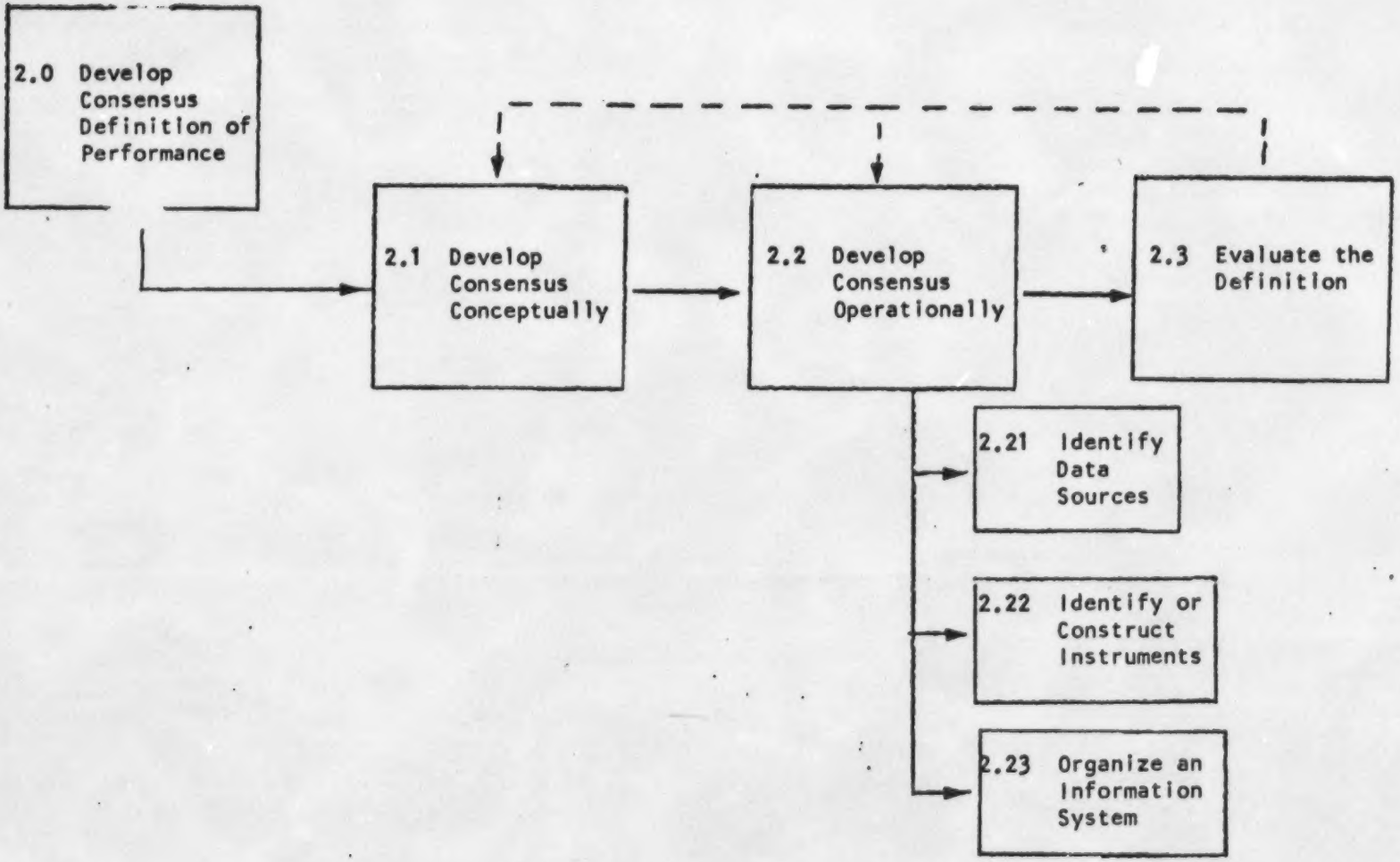
2.0 Develop Consensus Definition of Performance
2.1 Develop Consensus Conceptually
2.2 Develop Consensus Operationally
2.3 Evaluate the Definition
2.21 Identify Data Sources
2.22 Identify or Construct Instruments
2.23 Organize an Information System

Figure 15. "Develop Consensus" flow chart.

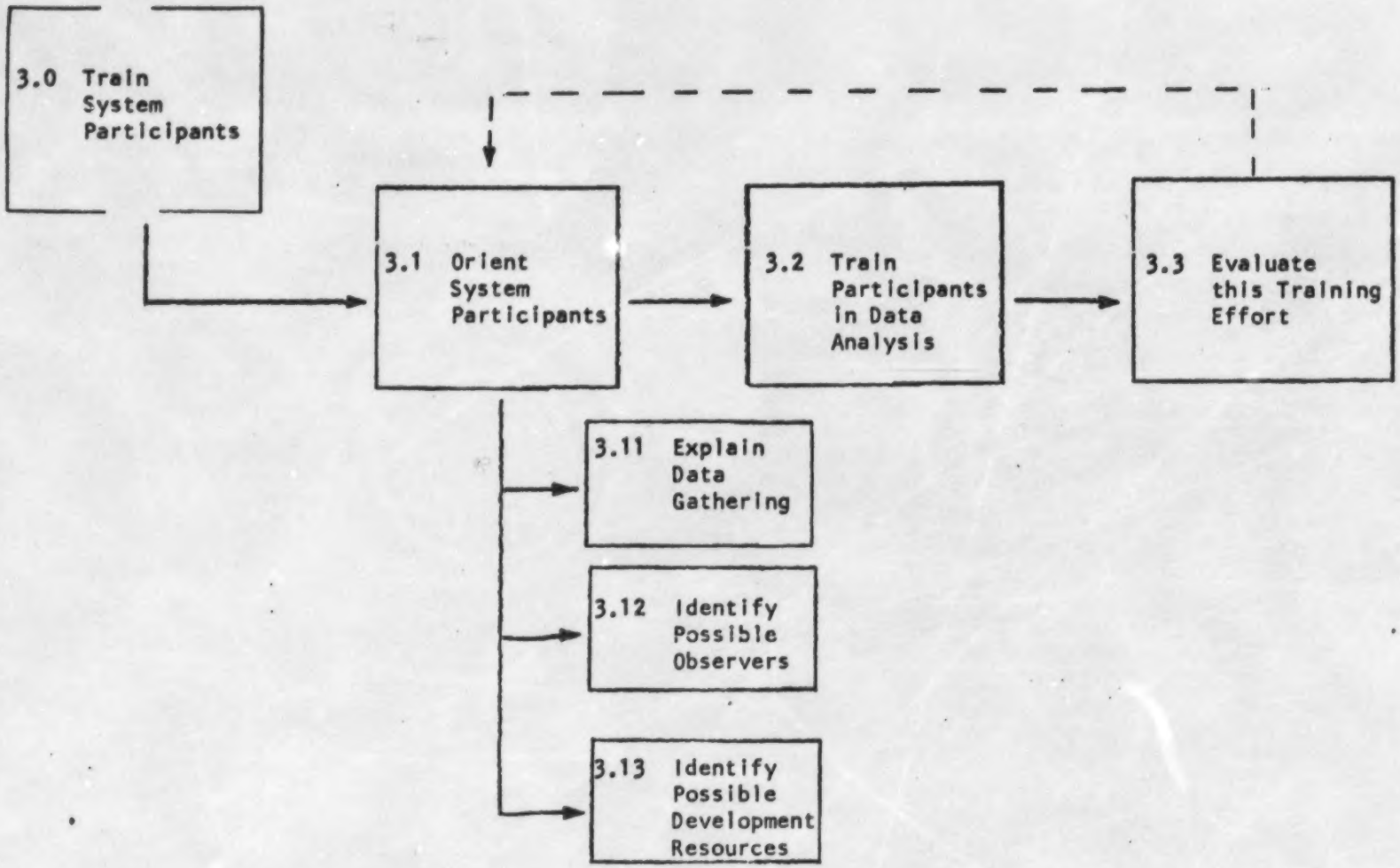
3.0 Train System Participants
3.1 Orient System Participants
3.2 Train Participants in Data Analysis
3.3 Evaluate this Training Effort
3.11 Explain Data Gathering
3.12 Identify Possible Observers
3.13 Identify Possible Development Resources

Figure 16. "Train System Participants" flow chart.

(4.0) Field-Test Systemic Appraisal

As figure 17 demonstrates, it would be necessary to field-test the entire systemic appraisal system by gathering subordinate data as a point of systemic linkage (4.1); producing developmental or diagnostic profiles (4.2) and evaluation profiles (4.3); and running the participants through both cycles (4.3 and 4.4). During implementation in some organizations, it might be advisable to phase in the cycles one at a time. While this author leans toward the development cycle as phase one of such a plan, arguments could be drawn either way as to whether evaluation or development would be first. The important thing to remember is that sooner or later both subsystems must be tested together for a true field test of the system.

(5.0) Evaluate the Field Test

Every systems approach features an evaluation component and this one is not an exception to that rule. Figure 18 presents a flow chart of functions associated with the evaluation of field test results. In order to accomplish system evaluation, this author would like to suggest the development of a "tissue committee" to replace implementation committees and monitor system operations. In hospital organizations, tissue committees examine organs removed in surgery as a way of monitoring surgical operations. This cross-check mechanism would question surgeons who extract healthy organs and suggest modification in the decision making process regarding decisions about initiating surgery.[34] Such committees made up of members representing various points of view could monitor the systemic appraisal process. A tissue committee for development might examine developmental programs and their effectiveness, selection patterns relevant to external observers and

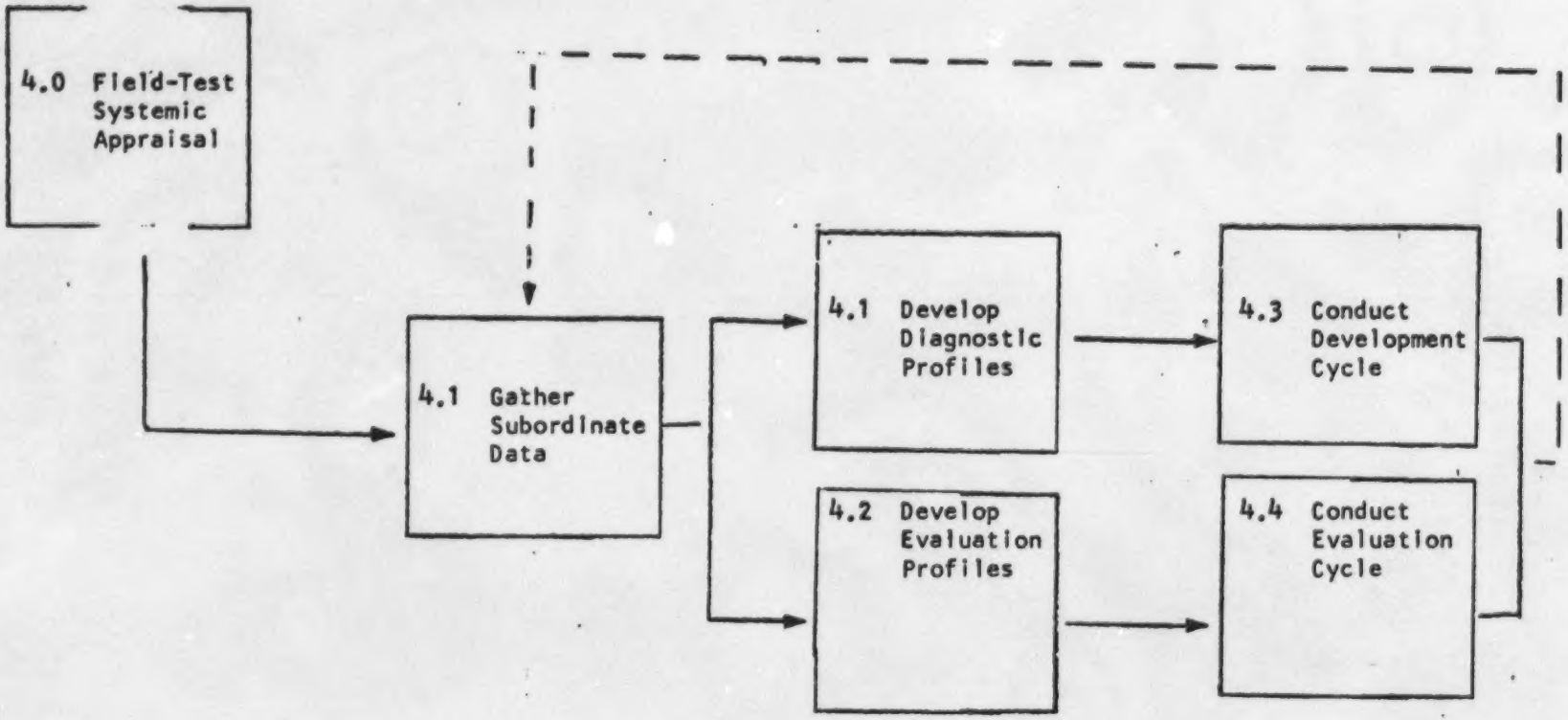
4.0 Field-Test Systemic Appraisal
4.1 Gather Subordinate Data
4.1 Develop Diagnostic Profiles
4.2 Develop Evaluation Profiles
4.3 Conduct Development Cycle
4.4 Conduct Evaluation Cycle

Figure 17. "Field Test Systemic Appraisal" flow chart.

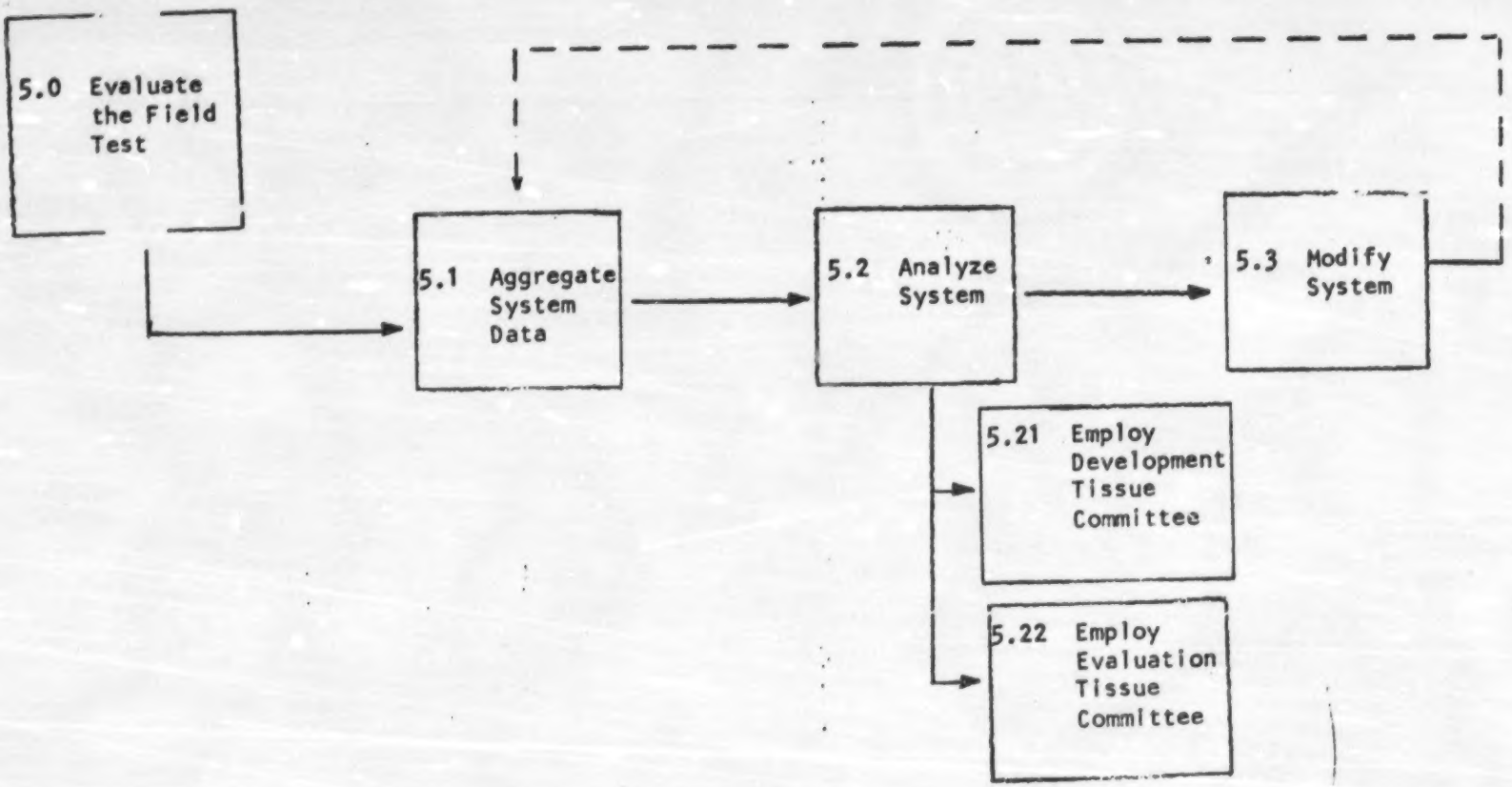
5.0 Evaluate the Field Test
5.1 Aggregate System Data
5.2 Analyze System
5.3 Modify System
5.21 Employ Development Tissue Committee
5.22 Employ Evaluation Tissue Committee

Figure 18. "Evaluate the Field Test" flow chart.

new resources for development. The evaluation tissue committee would examine evaluation profiles from subordinate ratings, external observations, and patterns appropriating organizational sanctions. These committees would necessarily operate in ways that protect the privacy of individual system participants. These committees would be responsible for aggregating the necessary data (5.1); analyzing the subsystems and the system as a whole (5.2, 5.21 and 5.22). Perhaps, meeting together annually or biannually in congress, these committees would make recommendations for the modification and improvement of the system (5.3). As the feedback loop notes, this process is repeated continually.

Conclusion

In beginning this project, the writer started with the straightforward notion that professional evaluation and development are functionally differentiated subsystems that have the potential for systemic intergration, i.e., professionals should work to improve in themselves that which the organization evaluates or organizations should help professionals develop those behaviors which the organization expects them to perform effectively. Appraisal was seen as the central vehicle by which professional leadership personnel can be developed and evaluated. The author's efforts were spurred by some encouraging past experiences linking the appraisal and development of school leadership personnel[35] and an invitation from the Mershon Center, The Ohio State University to study assessment and appraisal of leadership personnel in preparation for certain efforts related to the Center's on-going interest in leadership. After writing this paper, it became apparent that the notion, however simple in premise, would be a challenging one to operationalize in educational organizations.

The presentation of systemic appraisal was necessarily a sketch rather than a complete operational plan. There are two reasons for this type of presentation. First, the idea has little field test data to recommend its effectiveness.[36] Second, it is the feeling of this writer that such a system is best operationalized with the characteristics of a specific organizational setting in mind. Each organization can adapt and adopt this system in ways that make it viable given a particular situation. There is no way this model could be operationalized in this paper to meet the specific idiosyncrasies of particular educational organizations. Instead, an operational outline was presented along with a challenge to field test, evaluate and modify it in specific organizations. The operational details are probably best developed by system participants to meet organizational requirements.

Not only is this paper an outline of an idea and a challenge to experiment in various and differential organizational settings, but also a challenge to creatively adapt and modify system components and constructs. For instance, even though this paper emphasized subordinate ratings as part of the systemic link between development and evaluation subsystems, why not consider expanding the data base to include other reference groups like community members, organizational clients, or peers? Why should the system only focus on a leader's role expectations? Could the system provide for general, personal and professional enrichment? Of course, the possibilities are numerous and limited only by imagination.

Footnotes

1. H. D. Gutherie and D. J. Willower, "Ceremonial Congratulations: An Analysis of Principals' Observation Reports of Classroom Teaching," High School Journal, 56, (1972), 284-290.

2. S. Abramowitz, et.al., High School '77: A Survey of Public Secondary School Principals. Washington, D. C.: National Institute of Education, 1978.

3. R. A. Cooke and R. J. Coughan, "Developing Collective Decision-Making and Problem-Solving Structures in Schools," Group and Organizational Studies, 4, 1 (1979): 71-92 or J. W. Licata and E. C. Ellis, "Utilizing the Georgia Principal Assessment System as a Resource in Constructing an Instructional Component for School Administrators," CCBC Notebook (October, 1976), 5-15.

4. See The Beginning Teacher Project, The University of Georgia and the Georgia State Department of Education, 1975-1980 as an example.

5. C. R. Rogers, On Becoming a Person (Boston: Houston-Mifflin Company, 1968), Chapter 3, 39-58, "The Characteristics of a Helping Relationship."

6. J. S. Packard, The Norm of Teacher Autonomy/Equality: Measurement and Findings. Project MITT/Management Implications of Team Teaching, Center for Policy and Management, University of Oregon, Eugene, Oregon, 1967.

7. H. F. Wolcott, The Man in the Principal's Office (New York: Holt, Rinehart and Winston, 1973), 192-226.

8. Similar procedures have been employed by Licata and Ellis in Project ROME-FOCUS where principals took responsibility for assessment and development of their schools and themselves. See Licata and Ellis, "Utilizing the Georgia...".

9. C. D. Ellett, et.al., "The Relationship of Principal's Competencies and Meaningful School Outcome Measures: a Field Test of the Georgia Principal Assessment System," CCBC Notebook, 3, 5, 1976.

10. See C. D. Ellett and D. A. Payne, Principal Performance Description Survey (Athens, Georgia: Project ROME, 1977) for an example of an instrument measuring teacher perceptions of their principal's performance.

11. J. W. Licata and W. G. Hack, "School Administrator Grapevine Structure," Educational Administration Quarterly, in press.

12. These propositions have been modified from statements about evaluator characteristics noted in M. Webster, Jr. and B. Sobieszek, Sources of Self-Evaluation (New York: John Wiley and Sons, 1974), 1-64 and 156-166.

13. See J. W. Licata and R. L. Norman, "The Triangulation Interview Form," Education 99, 3 (1979), 321-325 for an explanation of this type of observation strategy.

14. Licata and Hack, "School Administrator ..."

15. T. Caplow, How to Run Any Organization (New York: Holt, Rinehart and Winston, 1976), 1-38.

16. See Vern Cunningham, "An Academy for Professional Development," Educational Administration, The Ohio State University, Columbus, Ohio, 1980 or J. W. Licata and E. C. Ellis, Project ROME-FOCUS Field Manual, ERIC Clearinghouse on Educational Management, 1980 for examples of developmental plans for educational leadership personnel.

17. D. A. Payne, et.al., "The Development and Validation of an Observation Instrument to Assess Competencies of School Principals," Educational and Psychological Measurement, 36 (1976), 945-952.

18. For instance, such procedures are explained in texts like R. A. Schmuck and P. J. Runkel, Handbook of Organizational Development in Schools (National Press, 1972), 115-116.

19. Schmuck and Runkel, 259-260.

20. C. D. Ellett and D. A. Payne, Principal Performance ... Figures 4 through 8 are in part taken from these documents and are used by permission of the authors.

21. C. D. Ellett, et.al., "The Relationship ..."

22. C. D. Ellett, The Continued Refinement and Development of the Georgia Principal Assessment System and Its Application to a Field-Based Training Program for Public School Principals, Project ROME Final Report, 1, Georgia State Department of Education, Thomas County Public Schools (Ga.) and College of Education, University of Georgia, 1976.

23. Ellett and Payne, Principal Performance ... See "Teacher Form," 6, Used by permiss on of authors.

24. Payne, et.al., 947.

25. Payne, et.al., 947-948.

26. Ellett and Payne, Principal Performance ... See "External Observer Form," 6, Used by permission of the authors.

27. Again, the profile forms used in figures 5-8 are in part taken from Ellett and Payne, Principal Performance ..., "Principal Form," 16-2 and "External Observer Form" 12. Used by permission of the authors.

28. Vern Cunningham, "An Academy ..." makes a similar suggestion.

29. M. Fullan, M. B. Miles and G. Taylor, "Organization Development in Schools: The State of the Art," Review of Educational Research, 50, 1 (Spring, 1980), 121-183.

30. See H. F. Wolcott, Teachers vs Technocrats (Eugene, Oregon: Center for Policy and Management, University of Oregon, 1977) for an in-depth description of such pitfalls, strategies and their consequences.

31. R. G. Havelock, The Change Agent's Guide to Innovation in Education (Englewood Cliffs, NJ: Educational Technology Publications, 1973).

32. W. G. Bennis, K. D. Benne and R. Chinn, The Planning of Change (New York: Holt, Rinehart and Winston, 1961).

33. R. A. Kaufman, Educational System Planning (Englewood Cliffs, NJ: Prentice-Hall, 1972), 75-105 presents a detailed explanation of function analysis and flow charting.

34. Tissue committees are noted by H. S. Becker, "Personal Change in Adult Life," Sociometry, published by the American Sociological Association (1964), 40-53 as an example a structural innovation in hospitals.

35. See. C. D. Ellett, The Continued ... and J. W. Licata, "An Internal Evaluation of a Field-Based Training Program for School Administrators," ERIC Clearinghouse on Tests, Measurements and Evaluation (1976) ED 135, 812.

36. While Ellett, The Continued ... and Licata "An Internal Evaluation..." present field test data on the effectiveness of linking assessment and development, little data is available about linkages with evaluation subsystems.